AD-A049 127 MIAMI UNIV OXFORD OHIO DEPT OF PSYCHOLOGY F/G 6/4
RESEARCH ON VISUAL PERCEPTION OF COMPLEX AND DYNAMIC IMAGERY.(U)
NOV 77 A J PANTLE F33615-76-C-5006

UNCLASSIFIED AMRL-TR-77-83 NL

1 of 2

ADA049127

12

AMRL-TR-77-83

RESEARCH ON VISUAL PERCEPTION OF COMPLEX AND DYNAMIC IMAGERY

ALLAN J. PANTLE
MIAMI UNIVERSITY
DEPARTMENT OF PSYCHOLOGY
OXFORD, OHIO 45056

NOVEMBER 1977

TECHNICAL REPORT AMRL-TR-77-83

Approved for public release; distribution unlimited.

AEROSPACE MEDICAL RESEARCH LABORATORY
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from Aerospace Medical Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with Defense Documentation Center should direct requests for copies of this report to:

Defense Documentation Center
Cameron Station
Alexandria, Virginia 22314

TECHNICAL REVIEW AND APPROVAL

AMRL-TR-77-83

This report has been reviewed by the Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Henning E. von Gierke

HENNING E. VON GIERKE
Director
Biodynamics and Bioengineering Division
Aerospace Medical Research Laboratory

AIR FORCE/56780/26 January 1978 — 150

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AMRL-TR-77-83
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): RESEARCH ON VISUAL PERCEPTION OF COMPLEX AND DYNAMIC IMAGERY.
5. TYPE OF REPORT & PERIOD COVERED: Final Technical Report, Sep 75 - Aug 77
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Allan J. Pantle
8. CONTRACT OR GRANT NUMBER(s): F33615-76-C-5006
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Miami University, Department of Psychology, Oxford, Ohio 45056
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62202F, 7233-05-18
11. CONTROLLING OFFICE NAME AND ADDRESS: Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio 45433
12. REPORT DATE: November 1977
13. NUMBER OF PAGES: 127
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 127p.
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Vision
Human Visual System
Spatial Frequencies
Sustained and Transient Spatial Frequency Channels
Visual Scan Paths
Motion Perception
Visual Responses to Dynamic Imagery

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report discusses three main research topics, each of which is interpreted in terms of a spatial frequency analysis model of the human visual system. (1) Parameters which affect the perception of clusters of line elements in dynamic, computer-generated displays are investigated. Transitions between perception of the moving cluster as a group and perception of the motion of the individual elements of the cluster are studied as a function of display frame rate, subject dark adaptation, image contrast, spatial orientation →

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

405 944

Block 20. Abstract (continued)

and various positional perturbations of elements in the cluster. The interaction of these parameters in interpreted within the framework of a visual system model consisting of sustained and transient spatial frequency channels. ~~(1)~~ The effect of spatial frequency components on eye movement in real-life scenes was studied by correlating eye movement patterns obtained from photographic scenes which were spatially high-pass filtered, low-pass filtered or unfiltered. Results indicate that areas of scenes containing strong low spatial frequency components attract attention more than areas containing strong high spatial frequency components. ~~(3)~~ Velocity coding in the human visual system was studied by means of moving sinusoidal gratings. Results with complex moving gratings indicate that velocity perception depends partially on the temporal variation of intensity at each point on the retina due to the moving pattern. A complex pattern containing a fundamental spatial frequency and higher harmonic components appears to be moving faster than a stimulus consisting of only the fundamental component and moving at the same velocity as the complex pattern.

This research indicates how interactions between sustained and transient channels can be used to characterize human visual response to dynamic displays or imagery. In particular, transient and sustained channels seem to be involved in the perceived organization of complex scenes; for example, the detection of targets against a cluttered background. The high and low spatial frequency tuning of the sustained and transient channels respectively appears to be a factor in determining eye movements and regions of interest in a complex scene. These studies provide a consistent framework on which to model the dynamic performance characteristics of the human visual system for complex image analysis.

Overview and Summary

There is considerable neurophysiological and psychophysical research which demonstrates the existence of separate sustained and transient channels in mammalian visual systems (e.g., Kulikowski and Tolhurst, 1973). Sustained mechanisms in humans have been shown psychophysically to be most sensitive to high spatial frequencies and relatively insensitive to movement or flicker. One characterization of the sustained channel that has proved successful in describing the results of experiments with static imagery is the Fourier-analyzer model (see Pantle for a summary, 1974). Transient mechanisms are more sensitive to low spatial frequencies and to movement or flicker. The temporal frequency response characteristics of the transient channel were explored in depth by Pantle (1975).

The major goals of the present research were to provide corroborative evidence for the existence of sustained and transient channels, to further specify their response characteristics, and to determine what contributions they make to the perception of complex and dynamic displays. In the section entitled "Perceptual Grouping and Movement of Line-elements in Dynamic Displays" we have described research with two different types of dynamic visual displays whose perceptual characteristics can, to some degree, be related to already known properties of sustained and transient channels. This is an important advance because it demonstrates that a whole area of research which describes the response of sustained and transient channels to sinusoidal patterns can be used to understand the findings in

another research area with a class of stimuli which are not sinusoidal gratings but which can be decomposed into their spatial frequency components. Moreover, new properties of sustained and transient channels can be inferred from the results with the dynamic displays. Experiments in the section entitled "Research on the Processing of Spatial Frequency Information in Complex, Real-life Scenes" were conceived within the framework of the Fourier-analyzer model described by Pantle (1974) for static imagery. Attempts were made to relate various aspects of human behavior--visual fixations, recognition, and informativeness judgments of real-life scenes--to the spatial frequency content of the scenes. Finally, in the section entitled "Research on Velocity Coding" we have reported the results of experiments which bear on the manner in which transient mechanisms might contribute to the judgments of the speed of moving patterns.

Contents

List of Illustrations

List of Tables

Perceptual Grouping and Movement of
Line-elements in Dynamic Displays

Research with two kinds of dynamic, line-element displays is reported in this section. In the displays employed in the element-group experiments, the elements which are of interest and which are perceived to move are presented against an otherwise blank (noise-free) background. In the displays employed in the cluster movement experiments, the elements which are of interest and which are perceived to move are presented against a cluttered background.

A Bi-stable Spatiotemporal Display

Pantle (1975) has described a psychophysical technique which can be used to isolate two separate visual mechanisms in humans. With this technique, sensations of stroboscopic movement are produced by a cyclic alternation of two stimulus frames. Frame 1 contains 3 black dots (*a*, *b*, *c*) arranged in a horizontal row on a white background. Frame 2 contains three identical dots (*d*, *e*, *f*), also arranged horizontally but shifted to the right such that the positions of dots *d* and *e* of Frame 2 overlap those of *b* and *c* respectively of Frame 1 (see Fig. 1). Depending upon stimulus conditions, the spatiotemporal display gives rise to either of two percepts: either an observer perceives a group of 3 dots moving *in toto* back and forth (*group movement*) or he perceives the overlapping dots of each frame as stationary and a third dot moving back and forth from one end of the display to the other (*element movement*). The relative proportion of element movement and group movement sensations reported by an

PRECEDING PAGE BLANK

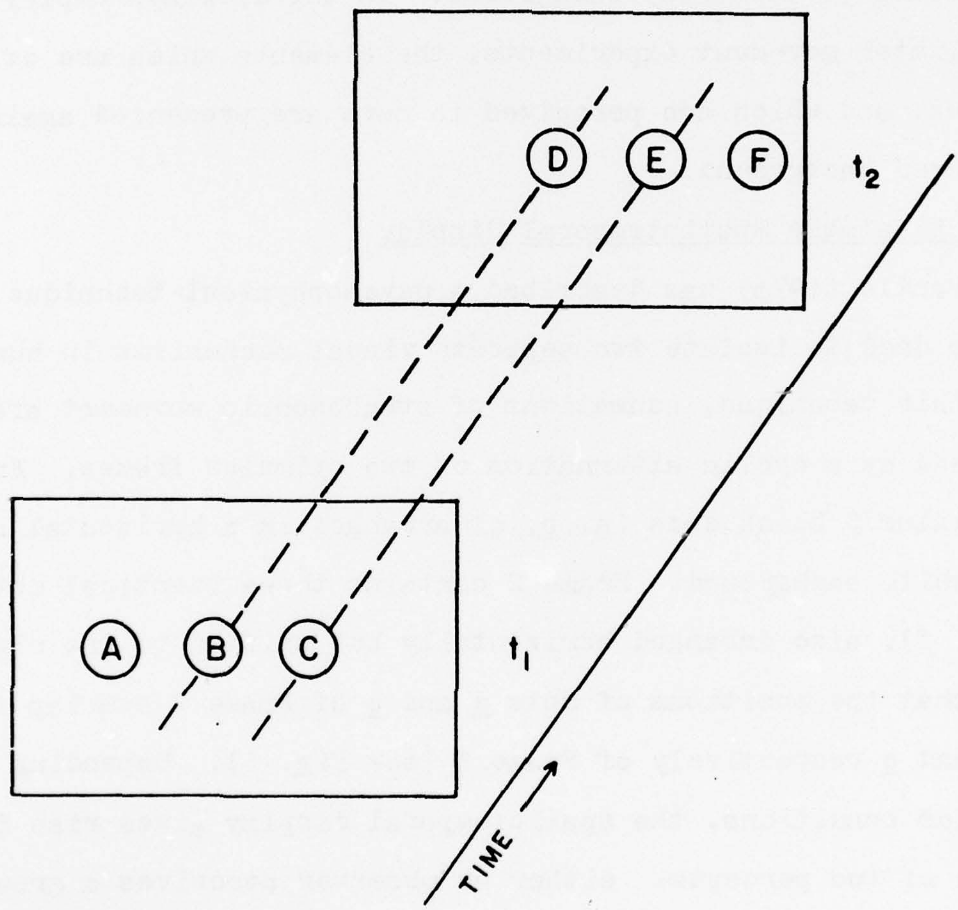


Fig. 1. A schematic representation of the spatiotemporal display used to create two competitive movement sensations.

observer with such a display in some way reflects the relative response (output) of two underlying visual mechanisms. That is, the relative proportion of element movement responses will increase whenever stimulus conditions favor the contribution of one visual mechanism; the relative proportion of group movement responses will increase whenever stimulus conditions favor the contribution of the second visual mechanism. By observing the effects of various changes of stimulus conditions on the relative proportion of element and group movement responses, it is possible to determine the response properties of the underlying mechanisms. Pantle (1975) has already used the element-group paradigm successfully and it can be exploited further in an attempt to gather more information about the response properties of the two visual mechanisms.

The effects of light intensity and contrast variables on the two visual mechanisms were studied in one series of experiments with the element-group movement paradigm. In a second series of experiments the effects of spatial variables on element and group movement sensations were investigated in order to determine whether the visual mechanisms underlying the element and group movement sensations were the sustained and transient channels, respectively.

Element-group Experiment I: ISI Illumination

When an observer views a pair of random-dot patterns which are presented successively and which each contain a region of dots which is the same except for a uniform displacement, the region appears to move back and forth (Anstis, 1970; Julesz, 1971; Bell and Lappin, 1973; Braddick, 1974). Braddick (1973) found that the perception of movement of the random-dot cluster could be weakened if the interval between the pair of successively presented patterns was illuminated with uniform light rather than kept dark. In Experiment I, we attempted to determine whether or not the presence of illumination in the interstimulus interval would differentially affect the mechanisms underlying the element and group movement sensations.

Method

Subjects. Nine undergraduate students and one graduate student served as subjects in the experiment. With the exception of the graduate student, all subjects served in partial fulfillment of course requirements. All ten subjects were naive about the purpose of the experiment and none had previous experience making psychophysical judgments.

Stimuli. Each member of the stimulus pair shown in Fig. 1 was presented in one channel of a three-channel tachistoscope (Gerbrands Harvard Tachistoscope; Timer Model 160A, Tachistoscope Switch Model 130S). Each stimulus frame was a white 4 in x 6 in unlined index card on which three dots were drawn in black ink. The relative positions of the dots on the two cards were the same as those in Fig. 1. The viewing distance was 81 cm, and at

this distance the diameter of each black dot was 40 min, with a center-to-center separation of 60 min visual angle between adjacent dots. The luminance of each black dot was 0.086 mL; that of the white background, 0.425 mL. The overall angle subtended by each stimulus frame was 9° horizontally and 6° 15' vertically.

In addition, a plain white card was inserted into the third field of the tachistoscope for use during light-filled ISI's. The third field could be adjusted to one of the following luminances: 0.00, 0.03, 0.09, 0.27, 0.80, or 2.40 mL.

Procedure. Each subject served in a single experimental session, the first few minutes of which were devoted to practicing the group and element movement judgments. Of the original 13 subjects who volunteered to serve in the experiment, three were dropped after the initial practice session because of confusion between the element and group movement sensations. No data were collected from these subjects.

During testing, each subject viewed 60 presentations of alternating stimulus frames (trials). ISI was held constant at 30 msecs and stimulus duration was held constant at 200 msecs. On any given trial the luminance of the blank field displayed during the ISI was randomly chosen from the six possible values under the constraint that no luminance would be repeated a third time until all other luminances had been presented twice. Each trial consisted of three complete cycles of the stimuli. For each trial the subject's task was to report whether he saw group or element movement.

Results and Discussion

The number of times that each subject reported group movement in each luminance condition was converted to a percent and entered into a Friedman analysis of variance by ranks. The pattern of results was consistent across subjects, and the effect of ISI luminance on the type of movement reported by the subjects was statistically significant ($\chi^2 = 35.89$, $\underline{df} = 5$, $\underline{p} < 0.001$).

The manner in which ISI affected the subjects' movement percept is shown in Fig. 2 where the mean percentage of group movement responses of all subjects is plotted as a function of the luminance of the uniform field during the ISI. The leftmost, circled point in the graph gives the mean percentage of group movement responses when the ISI was completely dark. In this condition the element movement sensation predominated, the subjects reporting group movement for only 27% of the stimulus sequences. The addition of light to the ISI had little or no effect on the type of movement seen as long as its intensity was less than the intensity of the stimulus frames (0.43 mL). The mean percentage of group movement responses for ISI luminances spanning a log unit (0.03 to 0.3 mL) remained approximately the same as that for a completely dark ISI. However, there was an abrupt change in the percentage of group movement sensations reported as ISI luminance was increased from 0.3 mL, a luminance lower than that of the stimulus frames, to 0.8 mL, a luminance higher than that of the stimulus frames. At the two highest ISI luminances, the group movement sensation predominated, the subjects on the average reporting group movement for 86-90% of

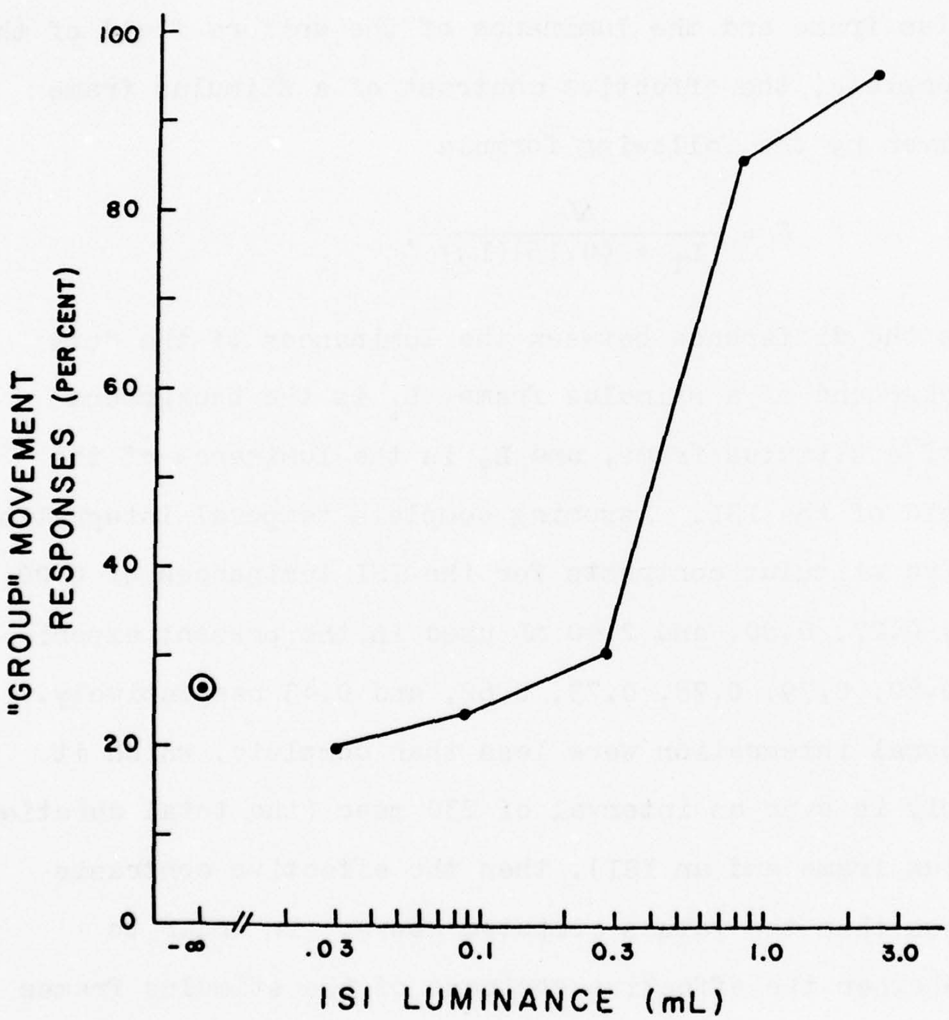


Fig. 2. Percentage of group movement responses as a function of the luminance of the interval between stimulus frames.

the stimulus sequences.

The influence of light in the ISI on the percentage of group movement sensations could possibly be due to a reduction in the effective contrast of the stimulus frames arising from temporal luminance summation. If the temporal integration of the luminance of a stimulus frame and the luminance of the uniform field of the ISI were complete, the effective contrast of a stimulus frame would be given by the following formula

$$\hat{c} = \frac{\Delta L}{L_1 + (0.15)(L_2)},$$

where ΔL is the difference between the luminances of the dots and the background of a stimulus frame, L_1 is the background luminance of a stimulus frame, and L_2 is the luminance of the uniform field of the ISI. Assuming complete temporal integration, the effective stimulus contrasts for the ISI luminances of 0.00, 0.03, 0.09, 0.27, 0.80, and 2.40 mL used in the present experiment were 0.80, 0.79, 0.78, 0.73, 0.62, and 0.43 respectively. If the temporal integration were less than complete, which it most probably is over an interval of 230 msec (the total duration of a stimulus frame and an ISI), then the effective contrasts would be less than the values computed above. In order to determine whether the effective contrast of the stimulus frames was the critical factor causing the changes in percentage of group movement sensations shown in Fig. 2, we directly manipulated the contrast of stimulus frames in another experiment.

Element-group Experiment II: Stimulus Contrast

Pantle (1975) has shown that a reversal of stimulus contrast from one stimulus frame to the next (from black dots on a white background to white dots on a black background) eliminates element movement sensations in the element-group movement paradigm. It is possible, therefore, that changes in the magnitude of stimulus contrast alone might alter the proportion of element and group movement sensations in the element-group movement paradigm. Such a finding would support the "change of effective contrast" interpretation of Experiment I. On the other hand, there is evidence from grating adaptation studies (see Sekuler, Pantle, and Levinson for summary, in press) that the response of at least one type of movement mechanism saturates at low stimulus contrasts (about 5X threshold) and is independent of contrast for a wide range of suprathreshold values.

Method

Subjects. Twelve undergraduate psychology students served in the experiment in partial fulfillment of course requirements. None of the subjects had previous experience making psychophysical judgments and all were naive about the purpose of the experiment.

Stimuli. Four pairs of stimuli were used, all with dot spatial arrangements identical to those used in Experiment I. However, each of the dots measured 9 mm in diam with a center-to-center separation between adjacent dots of 15 mm. The visual angle subtended by each dot was approximately 38 min. The center-to-center distance between adjacent dots was 1° 3' visual angle. The luminance of the background of each pair of stimuli

was 1.65 mL. The luminance of the dots for different pairs of stimulus frames was varied. For one pair, the luminance of the dots was 0.30 mL; for another pair, 0.77 mL; for the third and fourth pairs, 1.28 and 1.49 mL respectively. Thus, given a background of constant luminance, there were four pairs of stimulus frames with contrasts of 81.4, 53.3, 23.4, and 10.8 percent, where contrast is defined according to the following formula: $C = \Delta L/L$. ΔL is the difference between the luminances of the dots and the background, and L is the luminance of the background.

The stimuli were matte photographs of white index cards with dots drawn in black ink. Different exposure times were used to produce pairs of photographs (stimulus frames) with the different contrasts. Each pair of stimuli was presented with two channels of a three-channel tachistoscope.

Procedure. Subjects served in five experimental sessions conducted on separate days. During the first session subjects viewed alternating stimulus patterns with high contrast dots and described any movement sensations they experienced. Most subjects spontaneously reported element movement sensations when the ISI was short and group movement sensations when the ISI was long. The remainder of the first session was devoted to practice with various ISI's.

The first few minutes of each of the remaining four test sessions were devoted to practice with high contrast frames. Thereafter, during any particular test session, stimulus frames of a single chosen contrast were used for testing. During the

session there were 35 trials--five at each of seven ISI's (5, 10, 20, 30, 40, 50, or 70 msecs). Stimulus duration was held constant at 200 msecs and ISI was randomly selected with the constraint that no single ISI was repeated a third time until all others had been used twice. After each trial the subject's task was to report whether he had seen group or element movement.

Three separate 4 x 4 Latin squares were used to determine the order in which the four conditions of stimulus contrast (one per test session) were presented to the subjects. Each subject followed an order given by one row of one of the Latin squares. In this way the order of testing for different stimulus contrasts was counterbalanced across subjects. At the completion of the experiment, each subject had made five judgments for each combination of stimulus contrast and ISI.

Results and Discussion

Each function in Fig. 3 shows, for a single stimulus contrast, the mean percentage of group movement responses as a function of ISI. Each point on a curve is based on 60 responses, five for each of twelve subjects.

A repeated measures analysis of variance of the percentage of group movement responses demonstrated a significant main effect of ISI [$\underline{F}(6, 66) = 129.67$, $\underline{p} < 0.001$]. This result replicates earlier findings of Pantle (1975) that the percentage of group movement responses increases with an increase in the duration of the ISI. The present result extends the earlier findings in showing that the effect of ISI is present for low, as well as high, contrast stimuli. The main effect of stimulus

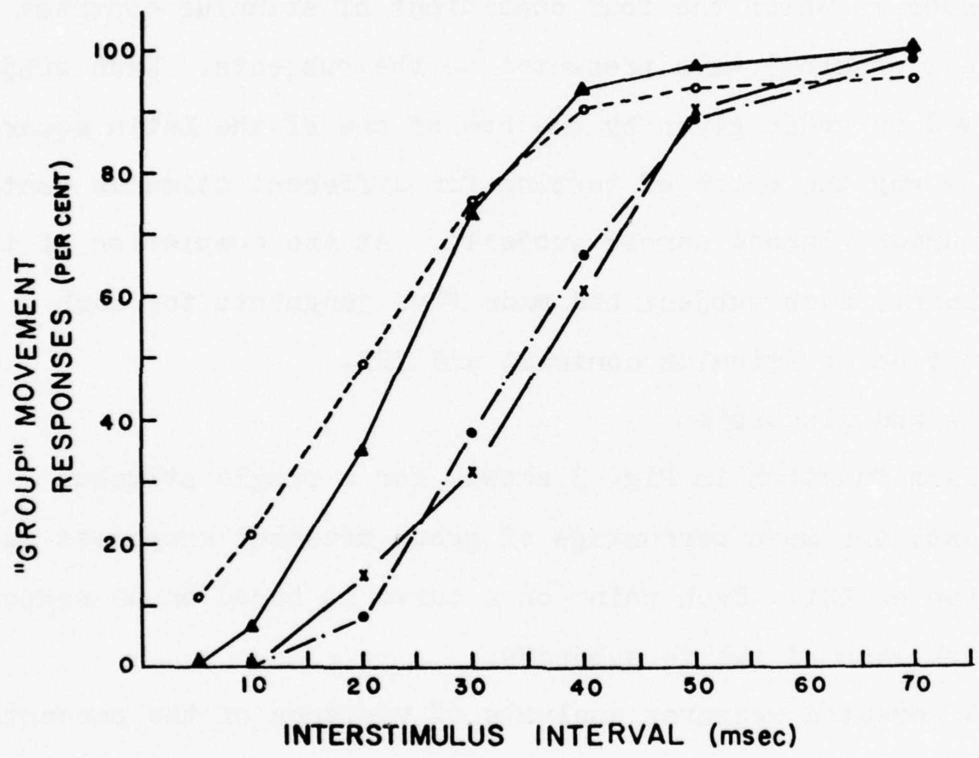


Fig. 3. Percentage of group movement responses as a function of the duration of the interstimulus interval. Stimuli of 10.8 percent contrast: o's; 23.4 percent contrast: ▲'s; 53.3 percent contrast: x's; 81.4 percent contrast: ●'s.

contrast was also significant [$F(3, 33) = 19.44$, $p < 0.001$]. Subjects reported a greater percentage of group movement sensations at lower contrasts. Except for the 70-msec ISI where the percentage of group movement responses is near 100 for all stimulus contrasts, points on the curves for 10.8 and 23.4 percent contrast fall above corresponding points on the curves for 53.3 and 81.4 percent contrast at all ISI's. The interaction of stimulus contrast and ISI was statistically significant [$F(18, 198) = 4.14$, $p < 0.001$], but the effect may be due entirely to the limitation of the range of the possible values of the dependent variable. Smaller differences between the curves for different contrasts occurred at the shortest (5 and 10 msec) and the longest (50 and 70 msec) ISI's, where the percentage of group movement responses approaches 0 and 100 respectively. Moreover, the shapes of the different contrast curves are almost identical, except for a horizontal displacement. As a measure of the displacement, we determined the ISI for which the element and group movement sensations were reported 50 percent of the time (hereafter called the transition ISI). This value was obtained for each subject for each stimulus contrast. For the four stimulus contrasts of 10.8, 23.4, 53.3, and 81.4 percent, the mean transition ISI's were 21.9, 25.0, 35.5, and 33.1 msecs respectively, and they correspond to the points where the four functions in Fig. 3 cross the 50 percent value of the ordinate. A repeated measures analysis of variance of the transition points revealed a significant effect of stimulus contrast [$F(3, 33) = 19.35$, $p < 0.001$]. The transition ISI's

were shorter for the two lower stimulus contrasts.

The changes in the percentage of group movement responses obtained in the present experiment with reductions of stimulus contrast are in the direction required to account for the results of Experiment I. However, the range of absolute stimulus contrasts over which reductions of stimulus contrast affected the percentage of group movement responses in Experiment II does not coincide with the range of effective stimulus contrasts over which reductions affected the percentage of group movement responses in Experiment I. In Experiment I, a change of effective stimulus contrast from 0.73 to 0.62 increased the percentage of group movement responses from 31 to 86 percent. In Experiment II, an even larger reduction of stimulus contrast from 81.4 to 53.3 percent produced no change in the percentage of group movement responses. The curves for these two contrast levels in Fig. 3 completely overlap. Not until the contrast was reduced to 23.4 or 10.8 percent, a few times threshold levels for the conditions of the experiment, was there any effect of stimulus contrast on the type of movement seen by the subject.

Two conclusions are warranted by the data: (1) At near threshold contrasts the mechanism underlying the group movement sensation more readily dominates the mechanism underlying the element movement sensation than it does at high contrasts; (2) some process other than temporal integration and reduction of effective stimulus contrast must be responsible for the effects of ISI illumination in Experiment I. The addition of light during the ISI increases the total amount of light received

by the eye during any given trial and during a test session. In Experiment III we investigated the effects of the level of light-dark adaptation on the proportion of element and group movement responses.

Element-group Experiment III: Light-dark Adaptation

It is generally held that, in humans, vision with high levels of illumination is mediated by cones and that vision with low levels of illumination is mediated by rods. This principle manifests itself during the course of dark adaptation after exposure to a bright flash of light. During dark adaptation there is a changeover from cone vision to rod vision. If there is a different input of rods and cones to the mechanisms mediating the element and group movement sensations, then one would expect a change in the proportion of element and group movement sensations produced by a constant spatiotemporal (element-group movement) display during dark adaptation. There are other reasons for expecting that the level of light-dark adaptation may influence element and group movement sensations. Physiological experiments show that the structure of receptive fields of visual neurons in the cat changes with the level of light-dark adaptation (Barlow, Fitzhugh, and Kuffler, 1957), and Breitmeyer (1973) has shown psychophysically that the temporal frequency response of movement-sensitive mechanisms studied with gratings shifts to lower frequencies as dark adaptation increases.

Method

Subjects. Five subjects, including the author (AP), participated in Experiment III. All subjects were experienced psychophysical observers.

Apparatus and stimuli. A three-channel Maxwellian-view system was used in the experiment. Two stimulus channels provided the two individual stimulus frames of the element-group movement

display, and the third channel provided a bright adaptation field.

The light beams of the two stimulus channels originated from a single source, a tungsten-filament bulb (#1133) run at 5.4V by a regulated power supply. In one of the stimulus channels, a collimated light beam was passed through a photographic transparency to produce one of the stimulus frames of the element-group movement display. The second stimulus channel provided the second stimulus frame. The retinal illuminance of the dots in each stimulus frame was approximately 110 trolands; the retinal illuminance of the background of each stimulus frame was 1100 trolands. Each dot subtended a visual angle of 53', and the center-to-center distance between a pair of adjacent dots was 1° 31'. Each stimulus frame as it appeared to the subject was circular with a diameter of 12° 37'.

The two stimulus frames were presented in alternate succession to the subject by means of an episcotister placed at a focal point in the optical paths of both stimulus channels. A variable-speed motor in combination with a multi-ratio speed reducer allowed the subject to control the speed of rotation of the episcotister, and thereby the rate of alternation of the stimulus frames. As the speed of rotation was increased, both the duration of each stimulus frame and the interval between them (the ISI, which was completely dark) decreased. Both of these decreases have been shown to produce a lower percentage of group movement responses in past work (Petersik, 1975). The ratio of stimulus duration to ISI duration remained the same for all rates of alternation of the stimulus frames.

The light beam of the third (adaptation) channel was supplied by a second tungsten-filament bulb (#1133) run at 6V by a second regulated power supply. The adaptation field provided by this channel was spatially uniform, was circular with a diameter of 12° 37', and had a retinal illuminance of 58,000 trolands. A stop was placed at a focal point in the adaptation channel to insure that the image of the source formed in the plane of the subject's pupil did not exceed the diameter of the subject's pupil.

All three channels were combined by beam splitters prior to the final lens of the optical system. The final lens formed three superimposed images of the two sources in the plane of the subject's right eye. A chin rest was used to stabilize the subject's head position. The retinal locations of each stimulus frame and of the adaptation field were concentric.

Procedure. The general plan of the experiment was to have each subject set the rate of alternation of the stimulus frames at a point which produced a multistable percept while viewing the element-group movement display at regular intervals following light adaptation. In general, the point of multistability was defined as that speed of rotation of the episcotister which yielded a sensation of spontaneous alternation between element and group movement sensations. At rates of alternation which were too slow the subject saw only group movement; at rates which were too fast, only element movement. The subject adjusted the alternation rate by changing the speed of the motor which controlled the speed of rotation of the episcotister in the two

stimulus channels. Prior to the start of the experiment all subjects spent 3-5 hours distributed over several different occasions familiarizing themselves with the apparatus and the method of setting their own multistable points. All subjects practiced setting their own multistable points until they were able to complete the adjustment within 10-15 secs.

Each subject served in three experimental sessions. At the beginning of each session, the subject dark adapted for approximately 2 mins. Next, the subject viewed the adaptation field alone with his right eye for a 90-sec interval. Following 90 secs of light adaptation, the experimenter occluded the adaptation field and the subject made his first multistable point setting. After finding his point of multistability, the subject turned away and faced a dark wall until the next time he was required to set his point of multistability. Exactly 90 secs after the <u>beginning</u> of the first adjustment of his point of multistability, the subject made a second adjustment. This sequence of events continued so that the subject set his point of multistability every 90 secs over a period of 30 mins of dark adaptation. Thus, during any single session, each subject made 21 multistable point settings.

Following each adjustment of the point of multistability by the subject, the experimenter recorded the speed of rotation of the episcotister and randomly reset its speed for the next trial. At the completion of the experiment each subject had made a total of 63 multistable point settings, three at each of the 21 delays after the adaptation field was extinguished.

Results and Discussion

The results of Experiment III are summarized in Fig. 4. The function gives the rate of alternation of the stimulus frames which the subject required to see the element and group movement sensations alternate spontaneously (the point of multistability) as a function of the time after the adaptation field was turned off. Since ISI and stimulus duration both changed when the subject varied alternation rate, the point of multistability is expressed both in terms of ISI (left-hand ordinate) and stimulus duration (right-hand ordinate). Each datum point is the mean of 15 estimates of the point of multistability, three for each of 5 subjects. As can be seen in the figure, the subjects had to increase the stimulus duration and ISI (the period of alternation of the stimulus frames) in order to maintain a multistable percept as time in the dark increased. An analysis of variance showed that the change in the point of multistability is statistically significant [$\underline{F}(20, 80) = 19.14$, $\underline{p} < 0.001$]. The change in the point of multistability is most easily understood in the following way. If the subjects did not change the alternation period as dark adaptation progressed, the element movement sensation became predominant; i.e., grew in strength relative to the group movement sensation. Since an increase of either stimulus duration or ISI favors the group movement sensation, the subjects increased the alternation period in order to preserve the balance between the group and element movement sensations as dark adaptation progressed.

It is not possible to draw any firm conclusions about the

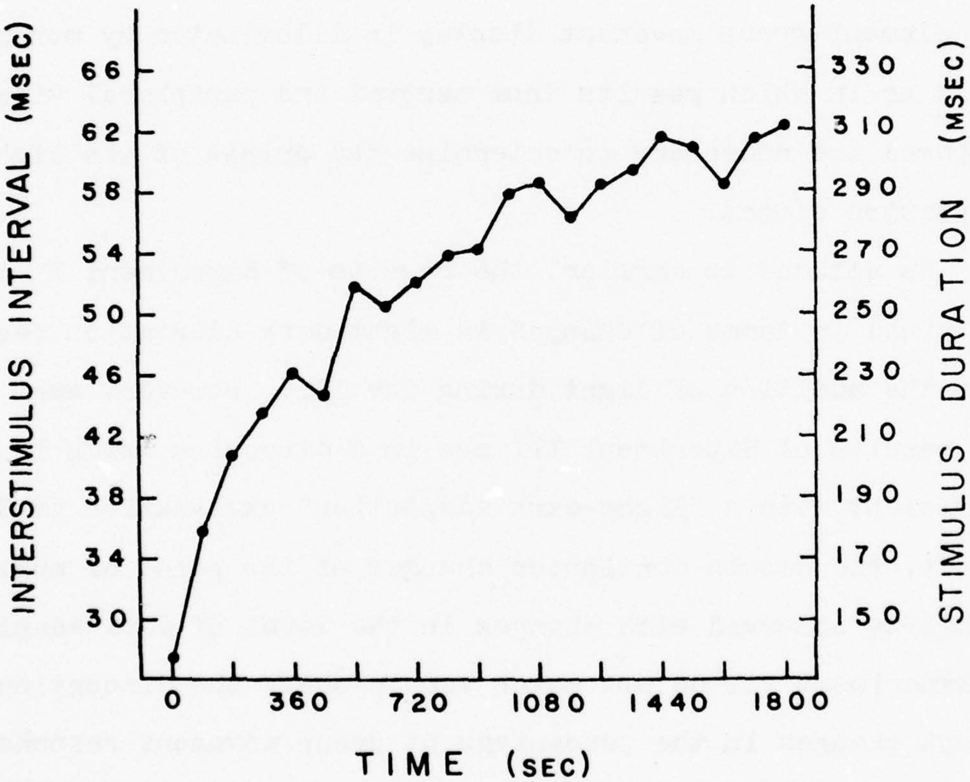


Fig. 4. The interstimulus interval and corresponding stimulus duration required to produce spontaneously alternating movement sensations as a function of dark adaptation time.

neural origin of the change in the point of multistability with dark adaptation. However, the lack of a distinct rod-cone break in the function indicates that it is more likely that the change of the multistable point is a result of a change in the response properties of the mechanisms which underlie the group and/or element movement sensations rather than a change from a cone- to a rod-dominated movement system. Further experiments in which the element-group movement display is illuminated by monochromatic light or in which results from central and peripheral viewing are compared are necessary to determine the origin of the light-dark adaptation effect.

As alluded to earlier, the results of Experiment I might be explained in terms of changes in light-dark adaptation resulting from the addition of light during the ISI. However, even though the results of Experiment III are in a direction which is consistent with a "light-dark adaptation" explanation of Experiment I, the smooth continuous changes of the point of multistability observed with changes in the level of dark adaptation in Experiment III do not match very closely the discontinuous abrupt changes in the percentage of group movement responses encountered in Experiment I as ISI illumination was changed. Apparently, the changes of ISI illumination in Experiment I were effective because of differences in rapid, temporal transients which accompany the ISI changes, and not because of differences in the overall level of light-dark adaptation. Experiments in which element and group movement responses are studied after various types of flicker adaptation may provide some further

insight into the results of Experiment I.

In summary, the three intensity and contrast variables studied in the foregoing experiments clearly affect the proportion of element and group movement sensations. While the exact manner in which the three variables alter the responses of the mechanisms which underlie the two sensations is not revealed by the experiments, comparisons among the experiments indicate that the variables are distinct; i.e., any one variable is not just a special case of one of the other variables.

Element-group Experiment IV: Orientation Perturbations

In the element-group experiments described thus far, the elements in each frame were dots. In the remaining element-group experiments the elements were vertical lines. Lines were used because they permit manipulation of spatial variables not possible with dots. In the first experiment with lines we introduced small systematic perturbations (variations) in the orientation of the two elements of one frame (elements which corresponded to d and e of the frame at time t_2 in Fig. 1) which overlapped two elements of the other frame. If the mechanisms underlying the element and group movement sensations are differentially sensitive to the changes of orientation, then one would expect the orientation perturbations to alter the proportion of element and group movement responses.

Method

Subjects. Eight undergraduate psychology students, six of whom served in partial fulfillment of course requirements, participated in the present experiment. Two of the subjects were experienced psychophysical observers, and all were naive with respect to the purpose of the experiment.

Stimuli. The stimuli were similar to those described in Experiment I and shown in Fig. 1, with the exception that the present stimuli were green vertical lines on a dark background instead of black dots on a white background. Each stimulus frame was presented on a CRT screen coated with a rapid decay phosphor (P31) by a PDP 11/10 computer. The viewing distance was 104 cm, and at this distance each line subtended 3' 40" visual angle in

width, and 3° 08' in height. The separation between adjacent lines was 1°. Stimulus frames were viewed with ambient room illumination, and the resultant luminance of the CRT screen was 0.12 mL. Against the background of the CRT screen, each of the stimulus lines generated by the computer had a luminance of 7.2 mL. The oscilloscope screen subtended 11° 45' vertically and 15° 15' horizontally.

In addition, the computer was programmed such that the left and center lines of the one frame in each pair (i.e., lines corresponding to elements _d_ and _e_ in Fig. 1) could be rotated about their center to either 0°, 3°, 6°, or 8° clockwise from vertical on any trial. The rotation is referred to as an orientation perturbation.

Procedure. Each subject served in five experimental sessions, the first few minutes of each being devoted to the practice of group and element movement judgments. During each experimental session, each subject viewed 40 presentations of the alternating stimulus frames (trials). During each session, the 40 trials were divided into two blocks of 20 trials. For any trial within a single block, an ISI of either 0, 17, 33, 50, or 67 msec was combined with one of the four orientation perturbations described above. All possible combinations of the five ISI's with the four orientation perturbations defined the experimental conditions. Within a block, the particular experimental condition employed on any trial was randomly determined with the constraint that no single ISI or orientation perturbation was presented on any three immediately successive trials.

On each trial, each subject viewed the alternating frames of a stimulus pair for 3 sec. Each frame in the sequence had a duration of 200 msec. For each trial, the subject's task was to report whether he saw group or element movement.

Results and Discussion

Each function in Fig. 5 shows, for a single orientation perturbation, the mean percentage of group movement responses as a function of ISI. Each point on a curve is based on 80 responses, ten for each of eight subjects. A repeated measures analysis of variance of the percentage of group movement responses obtained from each subject demonstrated a significant main effect of ISI [$\underline{F}(4, 28) = 89.78$, $\underline{p} < 0.0001$]. As in Experiment II, this result is a replication of earlier findings. An increase in the duration of the ISI produced an increase in the percentage of group movement responses. At short ISI's the element movement response was dominant, while at longer ISI's the group movement response was dominant. The main effect of orientation perturbation was also significant [$\underline{F}(3, 21) = 31.4$, $\underline{p} < 0.0001$]. Subjects reported a greater percentage of group movement sensations with greater perturbations of orientation. For example, as shown in Fig. 5, with an ISI of 33 msec the point on the 8° perturbation function falls above the corresponding point on the 6° perturbation function, which in turn is above the 3° perturbation function. The increase in the percentage of group movement responses (a decrease in the percentage of element movement responses) occurred only for short and intermediate ISI's because subjects already reported 100 percent group movement at long

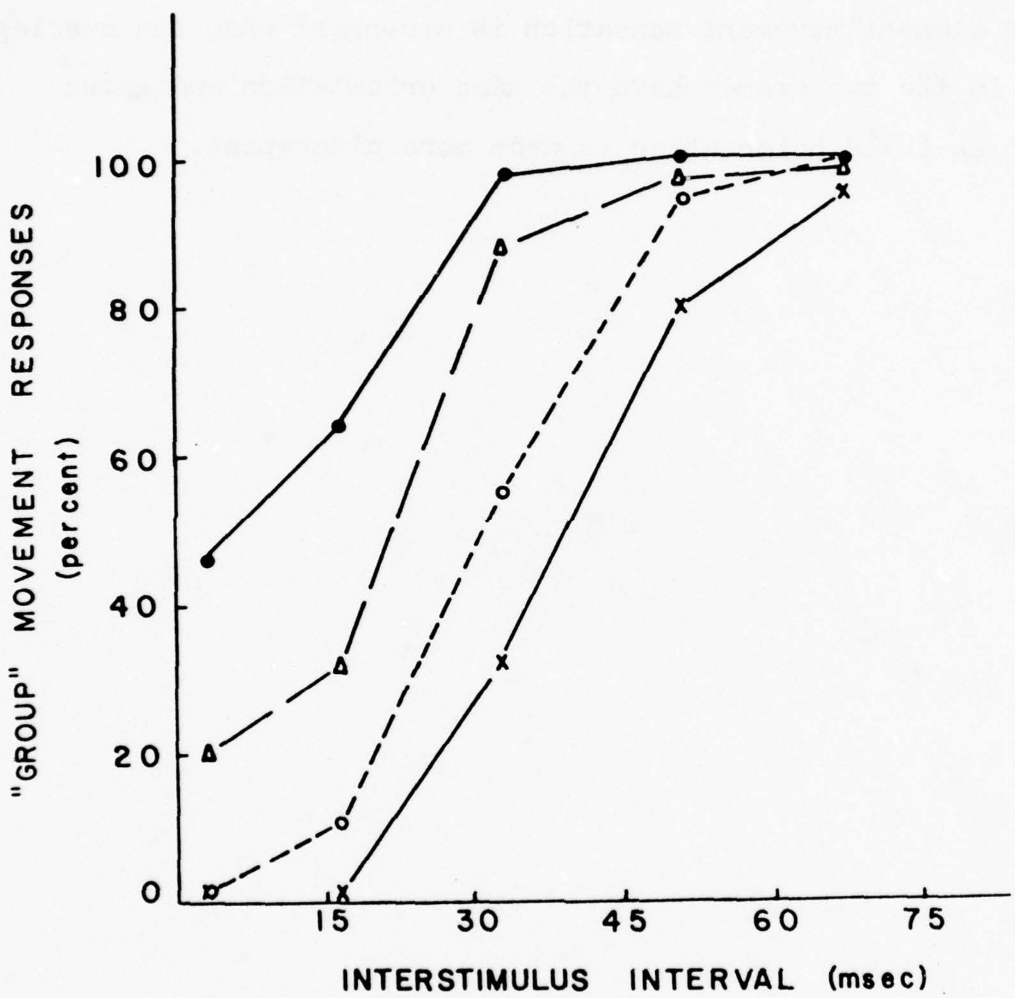


Fig. 5. Percentage of group movement responses as a function of the duration of the interstimulus interval. Orientation perturbation of 0° (control data): x's; 3°, o's; 6°, Δ's; 8°, •'s.

ISI's even with a 0° perturbation (the control condition). Consequently, the interaction between ISI and orientation perturbation is significant [$\underline{F}(12, 84) = 4.67$, $\underline{p} < 0.0001$].

To summarize, the response of the mechanism which gives rise to the element movement sensation is strongest when the overlapping lines in the two frames have the same orientation and grows weaker as their orientation is made more discrepant.

Element-group Experiment V: Position Perturbations and Element Spacing

Because the relative orientation of the overlapping lines in the element-group display was shown to be an important variable in the previous experiment, we decided to see whether or not slight perturbations in the relative positions of the overlapping lines would also affect the proportion of element and group movement responses. If the mechanism mediating the element movement sensation is sensitive to small changes (from one frame to the next) of the orientation of the overlapping lines, it might be equally sensitive to small changes in the positions of the overlapping lines. In a second experiment the positions of the lines corresponding to elements _d_ and _e_ in Fig. 1 were displaced by various amounts (hereafter called position perturbations) relative to their corresponding partners (_b_ and _c_ in Fig. 1) in the other frame. Because the position perturbations not only produce differences in the degree of correspondence of the overlapping lines, but also break up the spatial periodicity (regular spacing) of the lines in the frame in which the perturbations are introduced, a display which provided a control for aperiodicity was employed. In this display, the position of the element corresponding to _f_ in Fig. 1 was perturbed, making the elements of this frame aperiodic in space without changing the exact correspondence of the positions of the overlapping lines in the two frames.

In all previous element-group experiments, the spacing among the elements of each frame was held constant. However, it seemed

logical that the degree to which the lines were perceived as individual elements as opposed to a group would depend upon how closely they were spaced. Looked at in another way, the spatial frequency content of each frame is altered whenever the element spacing is changed, and the changes of spatial frequency content may differentially affect the mechanisms mediating the element and group movement sensations. Therefore, element spacing was also varied in the present experiment.

Method

<u>Subjects</u>. Six undergraduate psychology students, four of whom served in partial fulfillment of course requirements, participated in the present experiment. The two experienced psychophysical observers who served in Experiment IV also participated in the present experiment. All subjects were naive with respect to the purpose of the experiment.

<u>Stimuli and procedure</u>. Two stimulus frames which were identical to those described in Experiment IV and which contained no orientation perturbations (orientation perturbation = 0°) were used in the control conditions of the present experiment. The control stimuli were combined with each of five ISI's--0, 17, 33, 50, and 67 msec--to produce a total of five control conditions. As suggested above, it is possible that element spacing would also influence the relative percentage of group movement sensations reported at each ISI. Fig. 6A is a schematic representation of two stimulus frames with an element spacing <u>X</u>. In the present experiment, the spacing between adjacent lines in the control conditions measured 1°. The computer was programmed

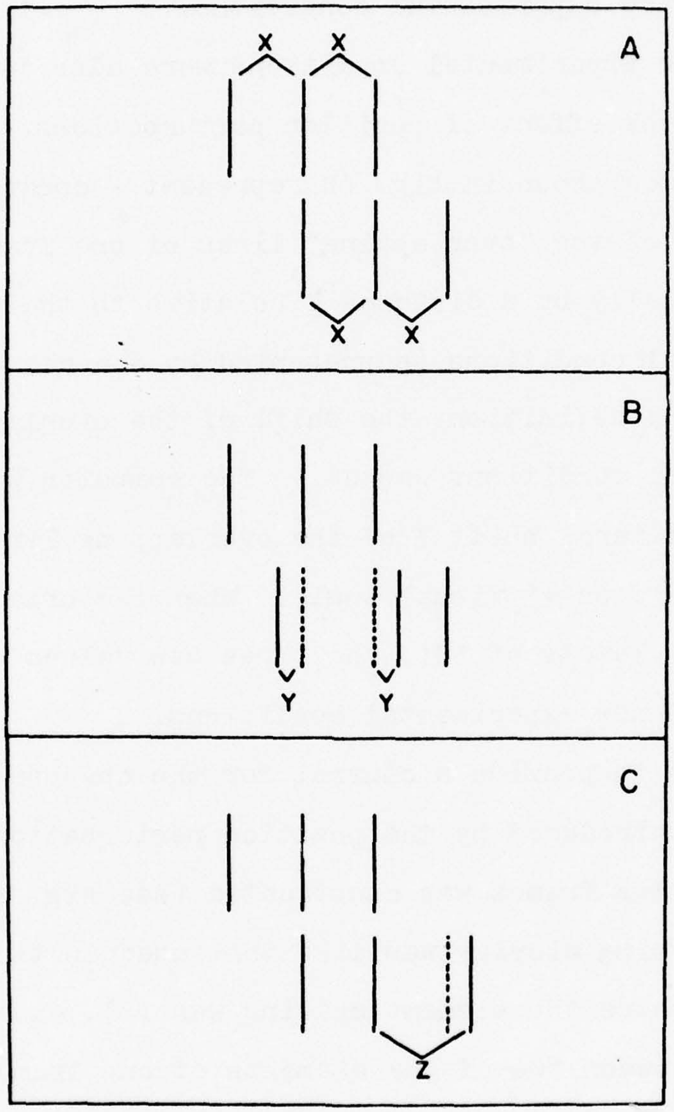


Fig. 6. Schematic representation of the variables of element spacing (X), position perturbation (Y), and aperiodic spacing (Z). See text for explanation.

so that the spacing X could also be 16', 8', or 4' visual angle. When factorially combined with the five levels of ISI (0, 17, 33, 50, and 67 msec), the three new values of X produced a total of 15 new experimental conditions.

Separate experimental conditions were also designed to investigate the effect of position perturbations. The two stimulus frames shown in Fig. 6B represent a condition in which the location of the "overlapping" lines of one frame have been shifted laterally by a distance Y relative to their locations in the control conditions (represented by the dashed lines in Fig. 6B). (By definition, the shift of the overlapping lines in the control conditions was 0°.) The computer was programmed so that the lateral shift Y of the overlapping lines could be either 16', 8', or 4' visual angle. When factorially combined with the five levels of ISI, the three new values of Y produced a total of 15 new experimental conditions.

In order to provide a control for the changes of the spatial periodicity introduced by the position perturbations, a separate pair of stimulus frames was constructed (see Fig. 6C). This aperiodic spacing display was like that used in the control conditions (where the element spacing was 1°), except that the distance Z between two of the elements of one frame was increased from 1° to 1° 40'.

The control, spacing, position perturbation, and aperiodic spacing conditions resulted in a total of 40 different conditions in the present experiment. Each subject served in ten experimental sessions, the first few minutes of each being devoted to the

practice of group and element movement judgments. During the remainder of each session, each subject viewed each of the 40 stimulus conditions (trials) described above. The order of presentation of the stimulus conditions was randomly determined. On each trial, the subject viewed the alternating stimulus frames for 3 sec. On each trial, the duration of an individual stimulus frame was 200 msec. For any single subject, the ten experimental sessions resulted in a total of 400 trials, ten for each condition.

Results and Discussion

Each of the three sets of conditions (element spacing, position perturbation, and aperiodic spacing) can be treated as a separate experiment. Figs. 7 through 9 summarize the results of the three sets of experimental conditions. The results of the control trials are shown in all three figures for comparison with the results of the experimental conditions.

Each function in Fig. 7 shows, for a single spacing $\underline{X}$ between adjacent stimulus lines, the mean percentage of group movement responses as a function of ISI. Each point on a curve is based on 60 responses, ten for each of six subjects. In general, the percentage of group movement responses increased as a function of ISI. A repeated measures analysis of variance of the percentage of group movement responses obtained from each subject verified the significant main effect of ISI [$\underline{F}(4, 20) = 19.97$, $\underline{p} < 0.0001$]. The main effect of spacing (magnitude of $\underline{X}$) was also significant [$\underline{F}(3, 15) = 9.05$, $\underline{p} < 0.002$]. Subjects reported a greater percentage of group movement sensations with

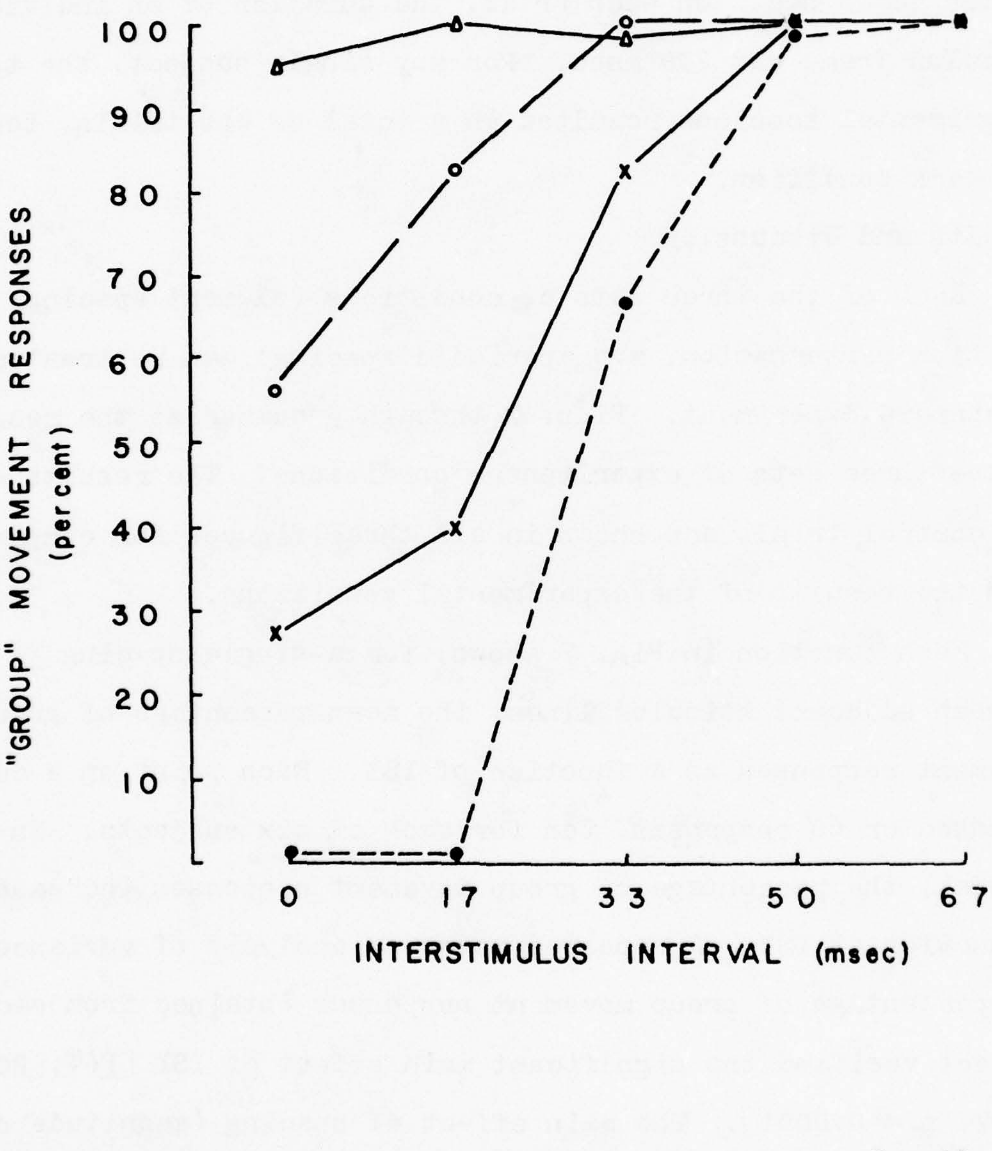


Fig. 7. Percentage of group movement responses as a function of the duration of the interstimulus interval. Position perturbation (Y) of 0 min (control data): •'s; 4 min, x's; 8 min, o's; 16 min, Δ's.

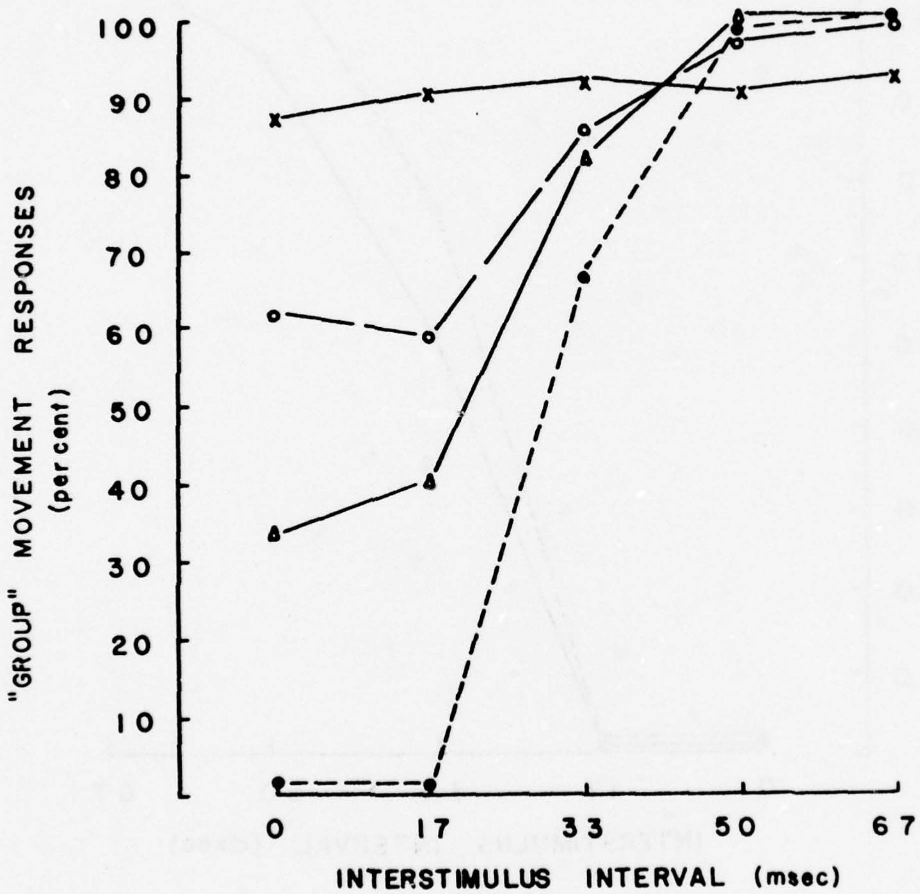


Fig. 8. Percentage of group movement responses as a function of the duration of the interstimulus interval. Element spacing (X) of 1 deg (control data): •'s; 16 min, Δ's; 8 min, o's; 4 min, x's.

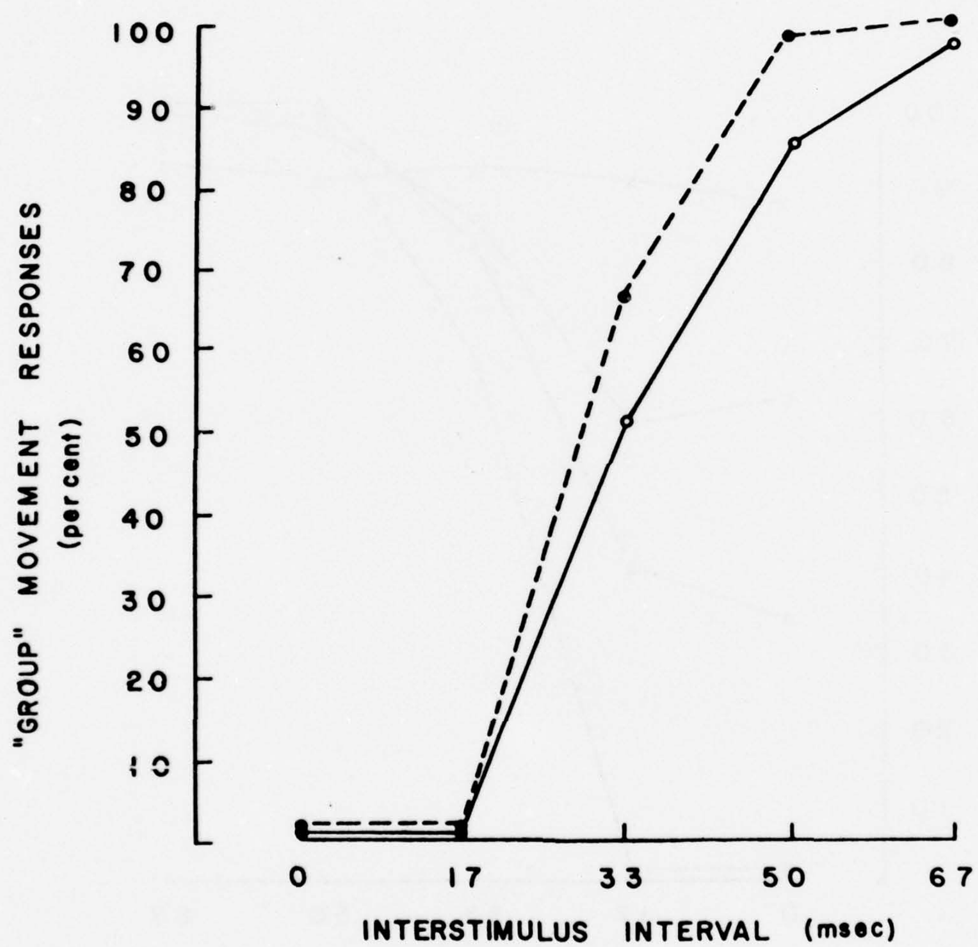


Fig. 9. Percentage of group movement responses as a function of the duration of the interstimulus interval. Periodic spacing (control data): •'s; aperiodic spacing: o's.

smaller spacings (smaller values of $\underline{X}$). For example, as shown in Fig. 7, for an ISI of 17 msec the point on the 4' spacing function lies above the corresponding point on the 8' spacing function, which in turn is above the 16' spacing function. The increase in the percentage of group movement responses occurred only for short and intermediate ISI's because subjects already reported group movement nearly 100 percent of the time at long ISI's even with the 1° spacing (control condition). Consequently, the interaction of ISI and spacing is also significant [$\underline{F}(12, 60) = 10.06$, $\underline{p} < 0.0001$].

Each function in Fig. 8 shows, for a single position perturbation, $\underline{Y}$, the mean percentage of group movement responses as a function of ISI. Each point on a curve is based on 60 responses, ten from each of six subjects. As in previous experiments, a repeated measures analysis of variance of the percentage of group movement responses obtained from each subject showed a significant main effect of ISI [$\underline{F}(4, 20) = 46.33$, $\underline{p} < 0.0001$]. In general, the larger the position perturbation the greater was the percentage of group movement responses. This trend was verified by a significant main effect of position perturbation (magnitude of $\underline{Y}$) [$\underline{F}(3, 15) = 34.38$, $\underline{p} < 0.0001$]. For example, the 16' position perturbation function, with the exception of one point, falls above the 8' perturbation function at the shorter ISI's. The 8' perturbation function, in turn, falls above the 4' perturbation function. The increase in the percentage of group movement responses occurred only for short and intermediate ISI's because subjects already reported group

movement on nearly 100 percent of the trials at long ISI's even with a zero position perturbation (the control condition). Consequently, the interaction between ISI and position perturbation is also significant [$F(12, 60) = 12.66$, $p < 0.0001$].

The dashed function in Fig. 9 represents the data for the control conditions in which Z (see Fig. 6C) was equal to 1°; the solid function, data for the aperiodic spacing condition in which Z was equal to 1° 40'. The control and aperiodic spacing functions are nearly alike. This means that the effect of introducing an aperiodicity into one of the frames of the display was minimal. At long ISI's the experimental function shows a somewhat lower percentage of group movement responses than does the control function. A repeated measures analysis of variance of the percentage of group movement responses obtained from each subject showed that the small effect of aperiodic spacing was statistically significant [$F(1, 5) = 10.14$, $p < 0.02$]. It is important to note that the small effect of aperiodic spacing is in a direction opposite that produced by position perturbations of the overlapping lines and cannot account for the effect of the position perturbations. As usual, the main effect of ISI was significant [$F(4, 20) = 44.01$, $p < 0.001$]. The interaction of ISI and aperiodic spacing was not significant [$F(4, 20) = 0.57$, $p > 0.05$].

The major results of the element-group experiments in which spatial variables were manipulated can be summarized as follows: (1) group movement sensations predominate over element movement sensations for long interstimulus intervals; (2) at short

interstimulus intervals, the element movement sensation predominates provided that the elements of each frame are spaced far enough apart and the overlapping lines of each frame are matched closely in position and orientation; otherwise the group movement sensation predominates; and (3) a break-up (within limits) of the spatial periodicity of the elements of one frame has a relatively small effect on the proportion of element and group movement sensations. In the next section we show how it is possible to account for the effects of the spatial variables under the assumption that the mechanisms underlying the element and group movement sensations are the sustained and transient visual channels, respectively. Given this interpretation, it is possible to use the results of the element-group experiments in which intensity and contrast variables were manipulated to specify yet unknown properties of the sustained and transient visual channels.

The Sustained-transient Channels Hypothesis as an Explanation of Element-group Experiments

The spatial variations of intensity in a horizontal direction across each of the stimulus frames (with a line-element spacing of 1°) of the element-group display can be analyzed into their Fourier components. Each frame has a high concentration of energy in a sinusoidal component whose half period corresponds to the distance between the end lines of each frame, i.e., a sinusoidal component of 0.25 c/deg. This component is due to the intensity of the lines acting as a group or "blob" on an otherwise dark background. Its spatial frequency is optimal for stimulating the transient visual channel. Each frame also has a large-amplitude component whose spatial period corresponds to the distance between adjacent lines in each frame (i.e., a sinusoidal component of 1 c/deg), as well as large-amplitude components whose spatial frequencies are whole number multiples of the 1 c/deg component. The spatial frequency components of 1 c/deg and greater are due to the individual lines, and their location in the spatial frequency spectrum is optimal for stimulating the sustained visual channel. When the frames are alternated at a slow rate with a relatively long ISI, the transient visual channel will respond strongly and pass the low spatial frequency component (0.25 c/deg) of each frame. Consequently, the three lines will be perceived as a group. On the other hand, when the frames are alternated at a fast rate with a short ISI, there is ample opportunity for the sustained mechanisms to integrate the stimulation produced by the overlapping lines of the two frames.

The overlapping, individual lines would be separated from the group and perceived as stationary because the sustained channels respond best to the high spatial frequency components of each frame (components with spatial frequencies $\geq$ 1 c/deg).

Next consider the effect of element spacing on the proportion of element and group movement responses. If the separation between the lines of each frame is decreased, frequency components whose spatial period derives from the decreased distance between adjacent lines become so high that they become less effective stimuli for the sustained mechanisms. For example, when the element spacing is reduced from 1° to 4', as in the experiment described earlier, the spatial frequencies of the components with large amplitudes which are due to the distance between adjacent lines will increase from 1, 2, 3, . . . c/deg to 15, 30, 45, . . . c/deg. The lines as a group, however, would still contain enough power at low spatial frequencies to stimulate the transient channel. Therefore, one would expect to obtain a greater proportion of group movement responses as element spacing is decreased, an effect which was observed. Furthermore, slight perturbations in the positions or orientations of the lines of one frame which overlapped the lines of the other would cause different high spatial frequency mechanisms (sustained mechanisms) to respond to the overlapping lines, but would have very little effect on the response of low spatial frequency mechanisms (transient mechanisms). If the sustained mechanisms which respond to the overlapping lines are not the same, then their effects would not be integrated and the peak response of the sustained

mechanisms would be lessened. Therefore, one would expect the orientation and position perturbations to increase the proportion of group movement responses, again an effect which was observed.

If it is accepted that sustained and transient channels underlie the element and group movement sensations, respectively, then the other element-group experiments with dots suggest that the sustained and transient channels have the following additional properties: Relative to the sustained channel's response, the transient channel's response to the spatial contrast of the elements of the frames of a spatiotemporal display increases (1) as the interval between the frames of the display is increasingly illuminated, (2) as the spatial contrast of the frames is reduced, or (3) as the level of light adaptation is raised.

It is possible to generalize the present results to other phenomena in which interactions between sustained and transient mechanisms are assumed to play a key role. For example, Breitmeyer and Ganz (1976) have suggested that metacontrast is the result of inhibition of the sustained channel's response to a test target by the transient channel's response to masking stimuli flanking the test target. On the basis of the present results it would be expected that, other things being equal, metacontrast would increase (1) as the interval between the test and masking stimuli is increasingly illuminated, (2) as the contrast of the test and masking stimuli are both reduced at near-threshold levels, or (3) as the level of light adaptation of the observer is raised. That there is considerable promise in

such generalized predictions is demonstrated by the fact that the results of the present experiment on position perturbations can already be used to explain a metacontrast experiment conducted by Stoper and Banffy (1977). Stoper and Banffy employed a display in which two vertical, flanking lines masked a central, vertical test line. The masking lines followed the test line in time. Stoper and Banffy found that the masking effect was substantially reduced or even eliminated if two preconditioning lines which were identical in shape and position to the flanking lines were presented before the test-mask sequence and remained on until the onset of the flanking (masking) lines. Presumably, sustained mechanisms integrated the stimulation produced by the preconditioning lines and the masking lines and inhibited the transient channel's response to the masking lines, a necessary requirement for metacontrast. In the same way that perturbations of the positions of the overlapping lines facilitated the transient channel's response to the frames of the element-group display, perturbations of the positions of the preconditioning lines in Stoper and Banffy's experiment would be expected to facilitate the transient channel's response to the masking lines and result in greater metacontrast. Stoper and Banffy employed such a manipulation and obtained the result predicted here.

Perception of a Moving Cluster of Elements on a Cluttered Background

Perceptual grouping of elements in a dynamic, visual display can be demonstrated to depend upon global form information. One example of a display that has been used to demonstrate the role of global form information is as follows:

Line segments are positioned randomly (0.5 density) in the cells of a 20 x 25 imaginary array (see Fig. 10). In a given pattern the elements (indicated by B's in Fig. 10) which are used to fill a small subset (3 x 10 cell area) of the entire array differ from background elements (indicated by A's in Fig. 10) in the remainder of the array. The A elements might be vertical line segments, and the B elements, horizontal line segments. A second pattern is generated in the same way except that the position of the 3 x 10 cell subset (cluster) is changed (see Fig. 11). When the two uncorrelated patterns of line segments are alternately presented at a rapid rate, subjects see the cluster of horizontal line segments moving *in toto* back and forth. Because the pairings of point-for-point intensities of the stimulus frames are random no matter how the two patterns are aligned with one another, and because there is no average brightness difference between the cluster and background areas of each frame, the perceived segregation and movement of the cluster must depend upon global form information present in the successively presented patterns. Some type of form or feature processing must occur at a site prior to or concurrent with the generation of the cluster movement signal.

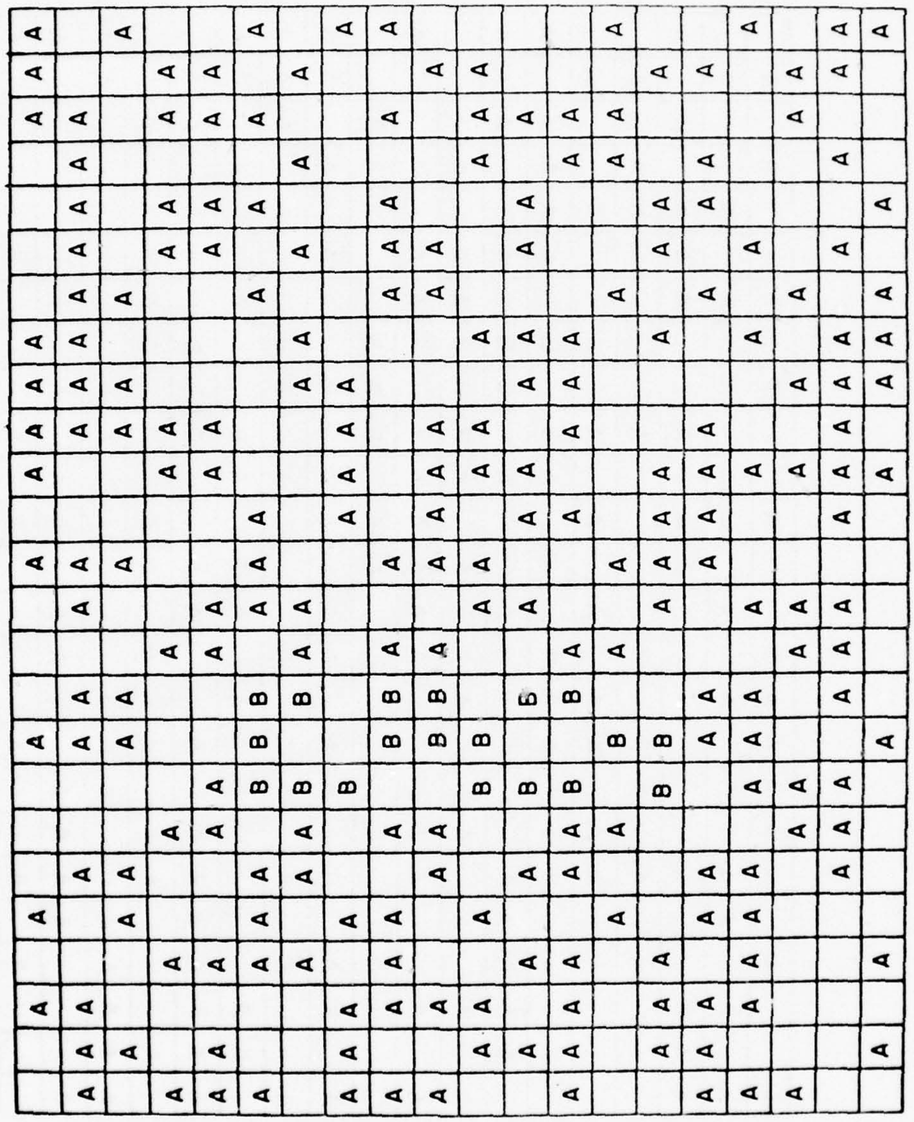


Fig. 10. A schematic representation of the spatial arrangement of the elements in one of two stimulus patterns required to generate global stroboscopic movement. See text for explanation.

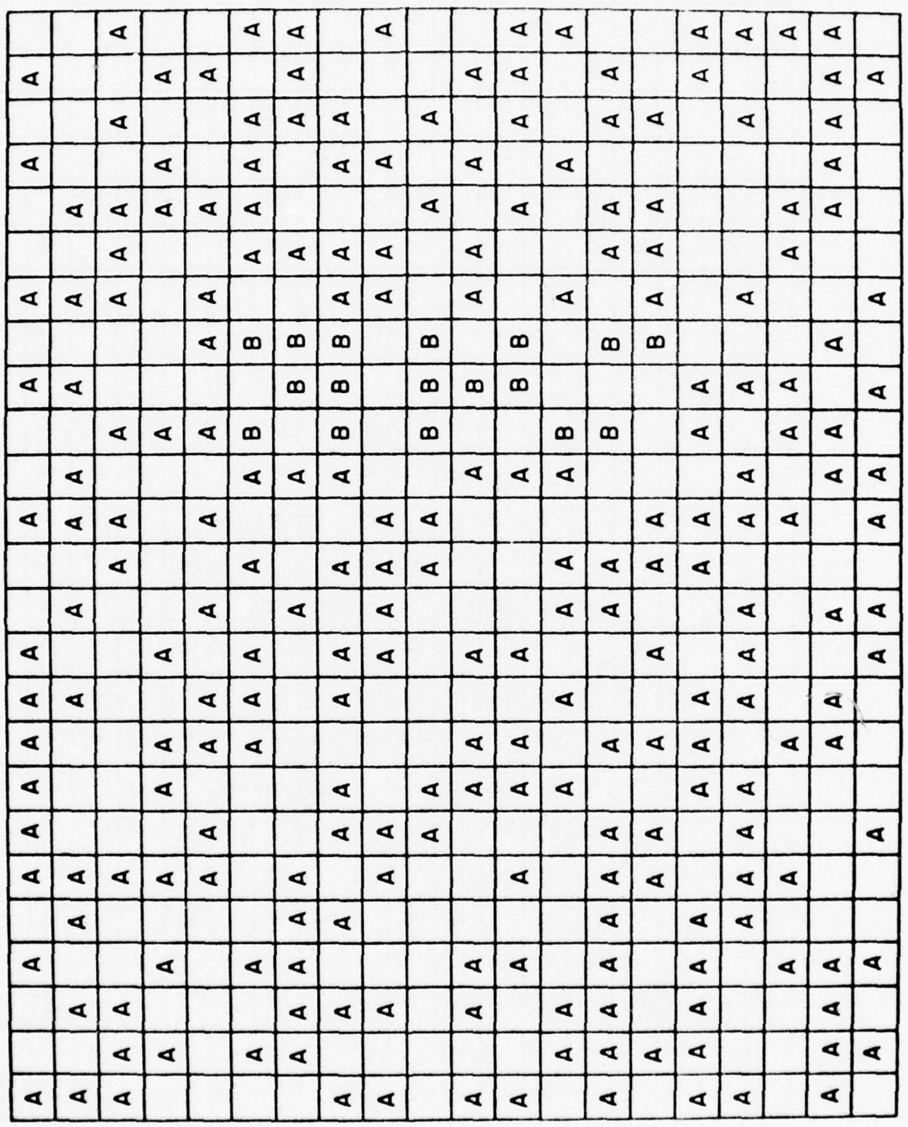


Fig. 11. A schematic representation of the spatial arrangement of the elements in the second of two stimulus patterns required to generate global stroboscopic movement. See text for explanation.

The apparent movement of the cluster disappears (adapts) if the display is viewed continuously and the frames are alternated at a rapid enough rate. When the movement sensation stops, the patterns appear only to flash "on" and "off," and oftentimes the cluster fragments; i.e., the elements comprising the cluster no longer appear to cohere together. The length of time that a particular sequence of frames can be viewed before the cluster movement sensation ceases (hereafter termed fragmentation time) was used as a dependent variable in the studies described in the following sections in order to gain some insight into the mechanisms responsible for the perceptual grouping of elements.

Cluster Movement Experiment I: Element Orientation

The psychophysical contrast threshold for a stationary sinusoidal grating varies as a function of the orientation of the grating; i.e., the threshold data with static gratings demonstrate that the human visual system, or at least the sustained visual channel, is anisotropic (see Pantle [1974] for a summary). The first cluster movement study was designed to determine whether the global form process responsible for cluster formation and movement was anisotropic.

Method

Subjects. Nine subjects, eight with normal vision and one optically corrected astigmat, served in the experiment in partial fulfillment of course requirements. All subjects were naive with respect to the purpose of the experiment, and none were experienced psychophysical observers.

Stimuli. All stimuli were generated in advance of the experiment and were stored on a disk by a PDP 11/10 computer. They were subsequently displayed on the face of an oscilloscope coated with a rapid-decay phosphor (P31) by the same computer.

Sequences of stimulus frames of the type described on page 52 were viewed binocularly and without a fixation point from a distance of 103 cm. At this distance one of the arrays (18 x 22) subtended a visual angle of 8° 14' vertically and 11° 6' horizontally. Individual line segments in each array subtended 3' 40" in width and 20' 3" in length. The cluster of differently oriented line segments subtended 4° vertically and 1° 20' horizontally; the distance between the edges of the

cluster in the two frames (see Figs. 10 and 11) was 3° 17'.

Three sets of stimuli were constructed for the present experiment. Each set was composed of two pairs of random-line patterns. Each pair of patterns was used for a different stimulus sequence, and it differed from the other pair only in terms of the particular positions of its randomly placed elements. Two pairs of patterns were constructed for each stimulus set so that any obtained differences between stimulus sets could not be attributed to the peculiarity of any unique array or pair of arrays. Half of the subjects in the present experiment always viewed one pair of frames from each set, while the remaining half of the subjects always viewed the alternative pair. Each of the sets of stimuli is described here: (1) The cluster was composed of vertically oriented line segments, while the background was composed of horizontally oriented line segments (VH condition). (2) A cluster composed of horizontal line segments was positioned on a background of vertically oriented line segments (HV condition). (3) The cluster was constructed from line segments oriented 45° counterclockwise from vertical, whereas the background was constructed from line segments oriented 45° clockwise from vertical (OB condition).

Procedure. Subjects served in one practice session and three experimental sessions, each conducted on a different day. During the practice session, each subject observed the apparent movement of the cluster that resulted from the alternation of the two stimulus frames of a pair. Additionally, each subject made some judgments about the perceptual fragmentation of the

cluster that typically occurs after prolonged viewing. Similar practice was conducted at the beginning of each experimental session.

Each experimental session consisted of 12 trials, which resulted from the combination of one pair of stimuli from each of the three stimulus sets described above with each of four frame durations (83, 117, 183, 267 msec). The ISI (67 msec) between the alternating frames remained constant throughout the experiment. During each session the order of trials was randomly determined with the constraint that neither the same stimulus pair nor the same frame duration was repeated on two successive trials.

On each individual trial the subject sat facing the oscilloscope screen with his right hand resting on the keyboard of a teletype connected to the computer. The computer was programmed so that, once a trial had begun, any signal from the teletype would terminate the trial and automatically record the trial's duration. On each trial, the subject viewed the alternating stimulus frames of a given condition until the apparent movement ceased or the cluster perceptually "fragmented." At this point the subject pressed a key on the teletype in order to terminate the trial. The duration of the trial recorded by the computer represented a single estimate of the subject's "fragmentation time" for that particular stimulus pair. That is, the duration recorded by the computer was the amount of time it took for the subject to report that the apparent movement of the cluster had ceased or that the cluster had perceptually fragmented.

Results and Discussion

The results of the optically corrected astigmat differed qualitatively from the results of the rest of the group. Therefore, her results are considered separately. For the eight normal observers, there appeared to be no systematic differences in fragmentation time between the two stimulus pairs which comprised each stimulus set. Therefore, the following results are collapsed across the two pairs within each stimulus set.

At the conclusion of the experiment, each subject had given three estimates of the cluster fragmentation time for each stimulus set and frame duration. From each set of three estimates yielded by each subject in each experimental condition, the median score was determined and taken as the estimate of each subject's fragmentation time. Each curve in Fig. 12 shows, for a single stimulus set, mean cluster fragmentation time as a function of frame duration. Each point on a curve is the average of the individual median estimates of cluster fragmentation time of the eight normal observers. For each function, fragmentation time increased with increases in frame duration. A repeated measures analysis of variance of the median fragmentation times demonstrated a significant main effect of frame duration [$F(3, 21) = 20.28$, $p < 0.0001$].

The main effect of stimulus set (HV, VH, and OB) was also significant [$F(2, 14) = 25.79$, $p < 0.0001$]. As shown in Fig. 12, for each frame duration fragmentation time was longest for the VH stimuli and shortest for the OB stimuli. The fragmentation times of the HV stimuli were shorter than, but relatively close

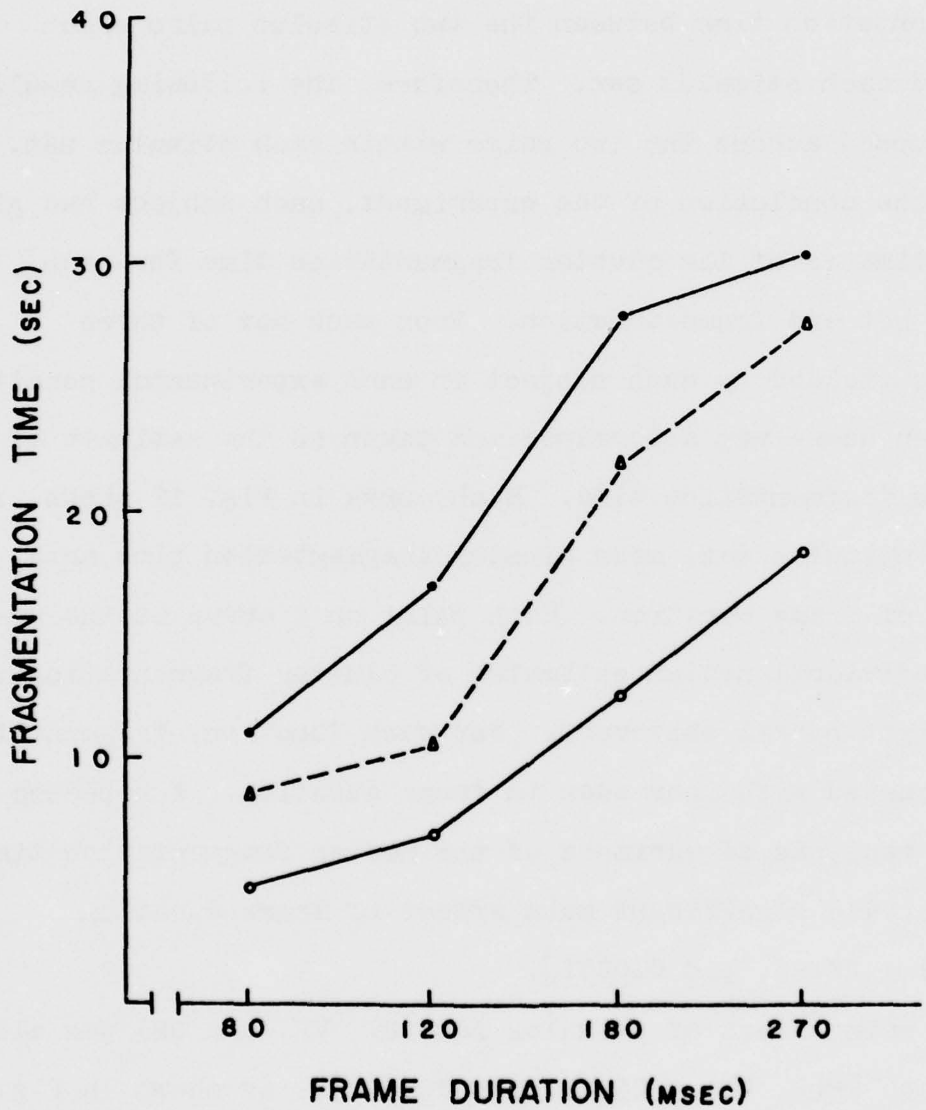


Fig. 12. Fragmentation time as a function of frame duration. VH frames: •'s; HV frames: Δ's; OB frames: o's. Normal observers.

to the fragmentation times of the VH stimuli. For every subject, the fragmentation time for OB stimuli was found to be shorter than that for the HV and VH stimuli. There was a significant interaction of stimulus set and frame duration [$F(6, 42) = 3.59$, $p < 0.01$].

It can be concluded that the form process responsible for cluster formation and movement is anisotropic. The shorter fragmentation times in the OB condition can be interpreted as evidence that the global form process is weak when it depends upon the processing of oblique elements. If this interpretation is correct, then the present results can be considered to be a simple extension of contrast sensitivity experiments which also demonstrate a weak response (low sensitivity) to oblique contours. While reduced sensitivity to oblique contours is a natural occurrence in observers with normal vision, it is not a natural occurrence in astigmatic observers. Indeed, Mitchell, Freeman, Millodot, and Haegerstrom (1973) found that adult observers with optically corrected astigmatism have low sensitivity for contours which would have been poorly focused in childhood due to their astigmatic condition. For this reason, an optically corrected astigmat, AMM, was included as a subject in the first cluster movement study. Due to her astigmatic condition in early childhood, horizontal contours would have been poorly focused on her retinae.

The results of the optically corrected astigmat, AMM, are shown in Fig. 13. Each point is the median of three separate estimates of cluster fragmentation time. The relationship between

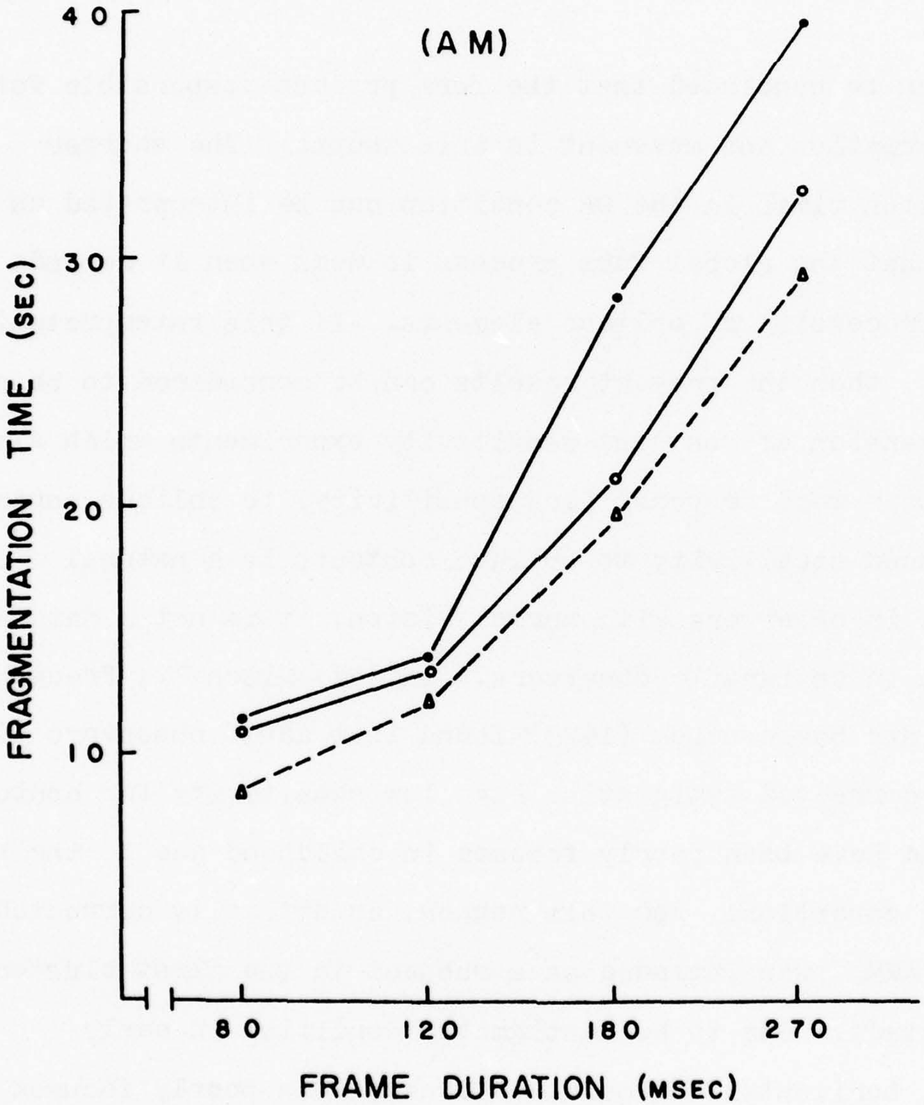


Fig. 13. Fragmentation time as a function of frame duration. VH frames: •'s; HV frames: Δ's; OB frames: o's. Astigmatic observer.

fragmentation time and element orientation for subject AMM was different from that of the normal observers (compare Figs. 12 and 13). For subject AMM fragmentation time for the OB stimuli was consistently longer than for the HV stimuli, and at short frame durations it was nearly equal to the fragmentation time of the VH stimuli. Unlike the normal observers, then, AMM's fragmentation times for the VH and HV conditions, which both require a response to horizontal contours, are not demonstrably longer than those in the OB condition. As with contrast sensitivity measurements, the pattern of anisotrophy in the fragmentation time data is related to optical defects present in childhood. The relationship is further evidence that the global form process is weak when it depends upon the processing of contours for which there is low sensitivity. In the next cluster movement study, we attempted to manipulate directly sensitivity to contours in various orientations to determine the effects of such a manipulation on fragmentation time.

Cluster Movement Experiment II: Selective Adaptation

The contrast threshold for a test grating can be elevated by prior adaptation to a grating in a similar orientation, but not by adaptation to a grating in a dissimilar orientation (see Pantle, 1974). For example, the contrast threshold for a vertical grating is raised by prior adaptation to a high-contrast vertical grating, but not by adaptation to a high-contrast oblique (45° from vertical) grating. In the present experiment, the fragmentation time for a cluster of horizontal elements on a background of vertical elements was measured after prior adaptation to an array of all vertical, all horizontal, or all oblique elements. If adaptation to the arrays is selective for orientation, and fragmentation time is related to sensitivity to contours in different orientations, then it would be expected that fragmentation times measured with the HV test display would be shortened by adaptation to either the array of all vertical or all horizontal elements, but not by adaptation to the array of all oblique elements.

Method

Subjects. Eight undergraduate psychology students participated in the experiment. Of these, seven subjects served in partial fulfillment of course requirements and were unpracticed observers. The eighth subject was a volunteer and an experienced psychophysical observer. All of the subjects were naive with respect to the purpose of the experiment.

Stimuli and apparatus. Since an adaptation-test paradigm was used in the experiment, sets of both adaptation and test

stimuli were constructed. Each of two test displays consisted of a pair of alternating HV frames as described in the previous cluster movement study. The reason for using two test displays was the same as in the previous cluster movement study. Each of the three sets of adaptation displays consisted of a pair of stimulus frames, and they are described below:

(1) Each frame was randomly filled (0.5 density) with horizontal line segments. No cluster of differently oriented line segments was present (all-H condition).

(2) Each frame was randomly filled (0.5 density) with vertical line segments. No cluster was present (all-V condition).

(3) Each frame was randomly filled (0.5 density) with oblique (45° clockwise from vertical) line segments. No cluster was present (all-O condition).

Viewing conditions, spatial subtense of each display and its elements, luminous intensity of the line segments, and the apparatus for presenting the test and adaptation displays were the same as in the last cluster movement study. Stimulus duration varied as a condition of the experiment, but ISI remained constant at 67 msec.

Procedure. Four of the eight subjects viewed only the first HV pair of test stimuli, while the other four subjects viewed only the second HV pair. Each subject served in four experimental sessions, the first of which was devoted to practicing the procedure of reporting the perceptual fragmentation of the apparently moving cluster with the HV stimuli. In addition, some practice trials were preceded by the presentation of one of

the pairs of alternating adaptation stimuli. Thereafter, the first few minutes of each experimental session were devoted to practicing the cluster fragmentation judgments under both adaptation and no-adaptation (control) conditions.

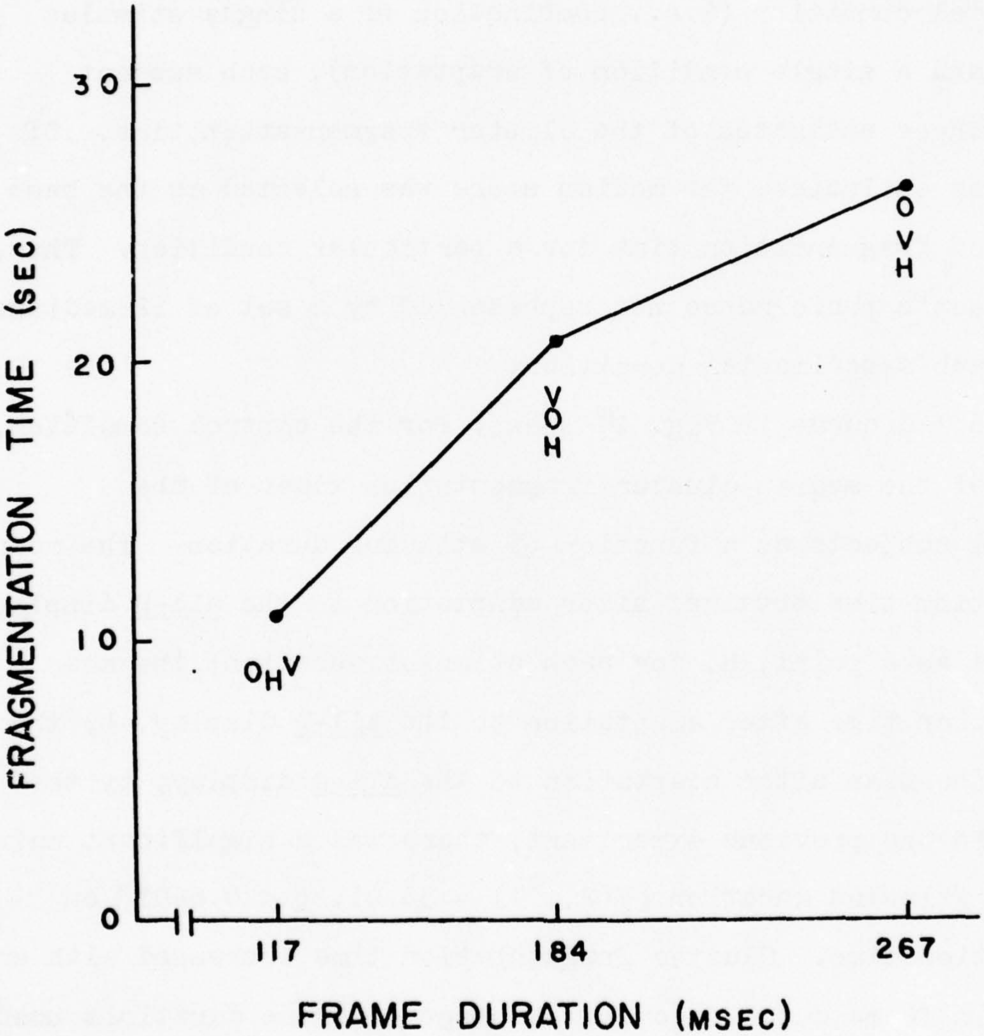
After the termination of the practice trials, each experimental session consisted of 12 trials which resulted from the combination of each of three frame durations (117, 183, and 267 msec) with each of four conditions of adaptation (all-H, all-V, all-O, and control [no adaptation]). Within a single experimental session, trials were randomly presented with the constraint that neither the same condition of adaptation nor the same stimulus duration was presented on any two successive trials.

On each control trial, a subject viewed the alternating HV display until perceptual fragmentation of the apparently moving cluster occurred, at which time he terminated the trial by pressing a key on the computer teletype. The computer recorded the trial duration (fragmentation time). On adaptation trials, each subject first viewed one of the three adaptation displays for 10 sec. After 10 sec, the adaptation frames were replaced without delay with the test display (alternating HV frames). As in control trials, the subject viewed the HV display until the cluster appeared to fragment or cease moving, at which point he pressed a key on the computer teletype. On the adaptation trials, the computer recorded the time that the HV display was presented before the subject pressed the key on the teletype (fragmentation time). Each subject completed three replications of the total set of 12 trials.

Results and Discussion

For thc eight subjects there were no systematic differences in fragmentation time between the two HV test displays. Therefore, the results were collapsed across the two HV displays. For each experimental condition (i.e., combination of a single stimulus duration and a single condition of adaptation), each subject provided three estimates of the cluster fragmentation time. Of these three estimates, the median score was selected as the best estimate of fragmentation time for a particular condition. Thus, each subject's performance was represented by a set of 12 medians--one for each experimental condition.

The solid curve in Fig. 14 shows, for the control condition, the mean of the median cluster fragmentation times of the individual subjects as a function of stimulus duration. The mean fragmentation time obtained after adaptation to the <u>all-H</u> display is plotted as a point, <u>H</u>, for each stimulus duration; the mean fragmentation time after adaptation to the <u>all-V</u> display, by the <u>V</u>'s; and the mean after adaptation to the <u>all-O</u> display, by the <u>O</u>'s. As in the previous experiment, there was a significant main effect of stimulus duration [$F(2, 14) = 36.01$, $p < 0.0001$] on fragmentation time. Cluster fragmentation time increased with an increase in frame duration over the range of frame durations used. The main effect of condition of adaptation was also significant [$F(3, 21) = 3.93$, $p < 0.05$]. The significant effect of condition of adaptation is due to the fact that the fragmentation times were consistently shorter on adaptation trials than on control trials. However, the size of the effect is small and not different



'ig. 14. Fragmentation time as a function of frame duration after selective adaptation to orientation. No-adaptation control data: •'s; all-V adaptation: V's; all-H adaptation: H's; all-O adaptation: o's.

for the different adaptation displays. The most interesting theoretical aspect of this result is the absence of a difference in fragmentation time in the three different conditions of adaptation. The interaction of frame duration and condition of adaptation was not significant [$F(6, 42) = 0.55$, $p > 0.05$].

The first cluster movement study suggests that relative sensitivity to the elements of a dynamic display is an important determinant of perceptual grouping. However, the second cluster movement study showed that, unlike contrast sensitivity measurements, the strength of perceptual grouping tendencies is not affected by selective adaptation to orientation.

More than relative sensitivity to elements must be involved in the strength of perceptual grouping tendencies with dynamic displays. At present it is not clear what the exact nature of the process is which integrates information about the elements into a whole or group. Nonetheless, from a theoretical standpoint, the cluster movement phenomena described in the last two experiments are interesting because it must be based on different mechanisms than the group movement sensation studied in the element-group experiments. In the element-group experiments the group movement sensation can be interpreted to be the result of a low-pass filter operation (by transient channels) which segregates the more intense group of lines (elements) from a darker background. As explained previously, however, the cluster movement sensation requires the processing of global form information, which is the only reliable cue for segregating the cluster elements from background elements. From an applied

standpoint, both types of grouping phenomena are important because they both are undoubtedly components involved in the perceived organization of complex scenes and in specific phenomena like camouflage.

Research on the Processing of Spatial Frequency Information in Complex, Real-life Scenes

Because the amount of information which is provided by the peripheral visual system exceeds the slow processing capacity of central mechanisms, it is important to investigate the strategies which are employed by the visual system in encoding complex scenes. Within the framework of the Fourier-analyzer model, such an investigation takes the form of evaluating the contributions made by different types of spatial frequency information to a variety of visual behaviors. In particular, in the following sections we attempted to assess the relative contributions made by the low and high spatial frequency components of a common set of visual scenes to recognition, judged informativeness, and visual scanning of the scenes.

Recognition of Complex, Real-life Scenes

Loftus and Bell (1975) have obtained some evidence which shows that the recognition of a visual scene depends upon two types of information: specific detail information and general visual information. The two types of information are acquired in different ways: for example, coding of specific detail information seems to depend upon a subject's visual scanning of a scene while coding of general visual information does not. In order to separate the effects of the two types of information on recognition performance, Loftus and Bell used photographs and line drawings as stimuli and relied on the reports of subjects about the information they used to recognize different stimuli. Within the framework of the Fourier-analyzer model, the two types

of information isolated by Loftus and Bell may be equivalent to the high and low spatial frequency components of the scenes. Viewed in these terms, manipulation of the two types of information can be reduced to spatial filtering, and the evaluation of the influence of each type of information on recognition performance can be more precisely defined than with the procedures employed by Loftus and Bell.

Pantle (1975) reported the results of an experiment in which subjects attempted to recognize spatially filtered, real-life scenes. During a study phase subjects viewed defocused versions of visual scenes (target set). During a test phase the target set was mixed with a set of additional scenes (null set) unfamiliar to the subjects, and they attempted to identify the target scenes. The performance of the subjects was better if the scenes presented during the test phase were defocused than if they were focused. High spatial frequency components which were present during the test phase, but not during the study phase, impaired recognition performance. While the finding is clear, its interpretation is not straightforward.

One possible interpretation of the results is simultaneous, perceptual masking. Pantle (1974) found that the threshold for a low-frequency (spatial) sinusoidal grating was elevated in the presence of high frequency background gratings. Harmon and Julesz (1973) demonstrated that the harmonic components in block portraits of human faces prevented recognition of the faces. In the experiment by Pantle (1975) on the recognition of real-life scenes, the presence of high spatial frequency components

during the test phase may prevent recognition of the scenes because they mask the low-frequency information studied by the subjects. The low-frequency information was the only information which was available during the study phase and upon which successful recognition performance could be based.

A second possible interpretation of the results is that the high-frequency components of the focused test slides add "noise" to the memory retrieval process. In this view the focused slides simply contain more information which is irrelevant to successful recognition performance than do the defocused test slides.

In order to decide between the perceptual masking and memory retrieval hypotheses, an additional recognition experiment was conducted. In this experiment the decrement in performance which resulted from the addition of low spatial frequency components during the test phase was compared to the decrement produced by the addition of high spatial frequency components. According to the perceptual masking hypothesis, the decrement produced by the addition of low-frequency components ought to be smaller than that produced by the high-frequency components. This prediction follows from the masking hypothesis because Pantle (1974) has shown that low spatial frequency gratings mask high spatial frequency gratings less than high frequency gratings mask low frequency gratings. The memory retrieval hypothesis makes no such prediction. There are no grounds for assuming that the memory retrieval process is disrupted more by irrelevant high spatial frequency information than it is by irrelevant low spatial frequency information or vice versa.

Method

Subjects. The subjects were 65 Miami University undergraduates enrolled in a lower-level psychology course. Participation in the experiment was a course requirement.

Materials and apparatus. The stimulus materials consisted of three sets of 20 slides: (a) unfiltered (U) slides were unmodified and focused normally during presentation, (b) low-pass (L) slides were unmodified and presented with the projector 5/12 of a turn out-of-focus, and (c) high-pass (H) slides were optically filtered leaving only high spatial frequencies and were presented with a focused projector (refer to Capell [1975] for a description of the filtering procedures).

The slides contained pictures of water (waterfalls, lakes, oceans, etc.), architectural structures (homes, large buildings, cities, etc.), and general outdoor scenes (fields, roadsides, etc.). The slides in each of the three sets were the same except for the mode of presentation: U, H, or L.

A slide projector (Kodak Carousel custom Model 800H) set at high luminance was used to present the stimuli. The stimulus duration was controlled by a millisecond timer (Gerbrands "300" Series Digital Model G1168) operating a shutter placed in front of the projector.

The subjects were seated in front of the screen so as to form an arc yielding an average viewing distance of 165 cm. At this distance, the stimuli subtended a visual angle of 45° 4' horizontally and 30° 52' vertically.

Procedure. A recognition-memory task was used and consisted

of two parts: a study phase and a test phase. During the study phase a subject viewed 10 slides one at a time for 200 msecs. The interstimulus interval was approximately 3 secs. The subject was instructed to look at the center of each slide, taking in as much of the picture as possible without concentrating on any particular portion of the slide. A 2-min retention interval followed the study phase.

During the test phase 20 slides were presented one at a time for 5 secs. The interstimulus interval was a blank white field which lasted 5 secs. Ten of the test slides were those presented during the study phase (target slides) and 10 were new slides (lures) of the same type of object or scene as the target slides.

The subject's task was to determine whether or not each slide was a target slide seen during the study phase. The subject recorded his response using a six-point confidence rating scale (---, --, -, +, ++, +++). The scale indicates both whether the subject thought the slide was a target and how confident he was in his judgment. For example, a "+++" response meant that the subject was very confident that the slide was a target and a "--" response meant that the slide was "probably new" or not a target slide.

Each of the subjects served in one of four experimental conditions: (a) low-pass slides presented during both study and test phases (LL), (b) low-pass slides presented during the study phase and unfiltered slides presented during the test phase (LU), (c) high-pass slides presented during both study and test phases (HH), and (d) high-pass slides presented during the study phase

and unfiltered slides presented during the test phase (HU). Each subject was assigned randomly to one of the four conditions. Since the subjects were not familiar with filtered slides, a brief explanation of what high- and low-pass filtered pictures looked like was given. Before each phase of the experiment the subjects were told which type of slide (U, L, or H) they would see.

The 10 target slides were chosen at random from the 20 slides available for use in each condition. The remaining 10 slides were used in the test phase as lures. The slides were presented in random order during the study and test phases, and the same orders were used in each of the four conditions. Each of the four conditions was run twice. The random order of slide presentation was different for the second presentation of the four conditions, and different subjects were used.

Results and Discussion

The dependent measure was a confidence rating made on a six-point scale (--- to +++) where "+++" meant that the subject judged the stimulus to be a target (a member of the study set) and that he was very confident of his judgment.

For each of the four experimental conditions, total response (confidence rating) frequencies were obtained for each of the six response categories (---, --, -, +, ++, +++) by summing frequencies across subjects and stimuli. The total frequencies for each response type and condition are given in Table 1. For example, in the HU condition there were 13 responses to lures in the "+++" category; that is, there were 13 instances where the subjects

Table 1

Absolute Frequencies and Cumulative Proportions of Confidence Ratings for Targets and Lures for the Four Experimental Conditions

Condition	Item type	---	--	-	+	++	+++	Condition	Item type	---	--	-	+	++	+++
		Frequencies								Frequencies					
HU (_N_=19)	Lure	22	33	48	24	50	13	LU (_N_=17)	Lure	31	36	22	30	36	15
	Target	22	33	26	48	35	26		Target	22	27	28	24	39	30
		Proportions								Proportions					
	Lure	1.00	.88	.71	.46	.33	.07		Lure	1.00	.82	.61	.48	.30	.09
	Target	1.00	.88	.71	.57	.32	.14		Target	1.00	.87	.71	.55	.41	.18
		Frequencies								Frequencies					
HH (_N_=13)	Lure	23	30	28	17	27	5	LL (_N_=16)	Lure	27	47	24	23	21	8
	Target	11	16	22	23	36	22		Target	12	37	22	22	33	34
		Proportions								Proportions					
	Lure	1.00	.82	.59	.38	.25	.04		Lure	1.00	.83	.54	.39	.24	.05
	Target	1.00	.92	.79	.62	.45	.17		Target	1.00	.93	.69	.56	.42	.21

thought they had previously seen a stimulus with a high degree of confidence when in fact it was not in the original study set (i.e., was not a target slide).

In order to plot a receiver operating characteristic (ROC) curve (Green and Swets, 1966), using confidence ratings, it is necessary to calculate the cumulative proportion of responses in each of the six confidence-rating categories. For each category the cumulative proportion was obtained by summing the frequencies for a given category and all categories above it (e.g., the cumulative proportion for category "+" would include the frequencies in categories "++" and "+++"), and then dividing by the total number of frequencies in all categories (see Table 1). For example, in the HU condition 57% of the responses to targets were in categories "+," "++," and "+++," whereas 46% of the responses to lures fell in the same categories.

Considering each entry for targets as analogous to the proportion of "hits" in a signal detection experiment, and each entry for the lures as "false alarms," there are five points which can be used to plot an ROC curve for each condition. The curves for each of the four conditions are shown in Fig. 15. The distance of each curve from the diagonal is a measure of recognition performance which is independent of the changes in the subject's response criterion. The farther a curve is from the diagonal, the greater was the ability of the observer to discriminate between targets and lures. As an example, it can be seen from Fig. 15 that the subjects in the HH condition found it easier to distinguish targets and lures than subjects in

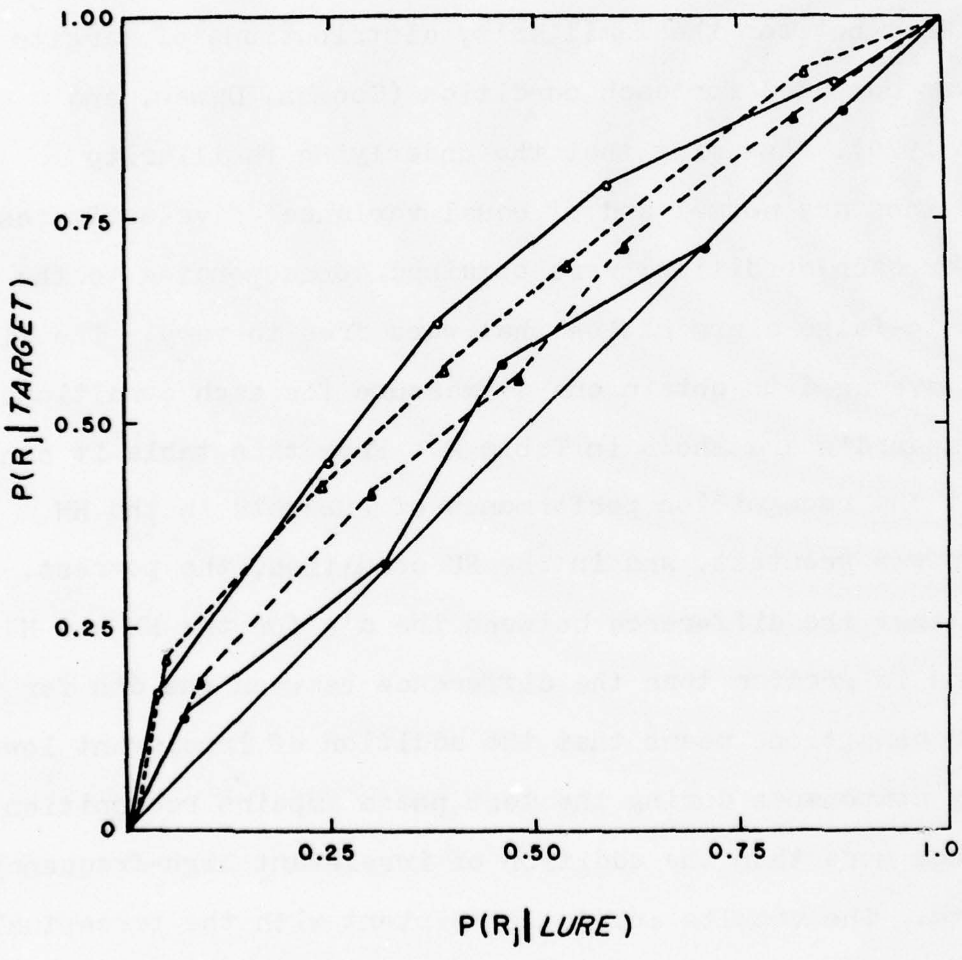


Fig. 15. ROC curves for the recognition of spatially filtered visual scenes. Condition HH: ●'s; condition HU: o's; condition LL: Δ's; condition LU: ▲'s.

the HU condition.

The distance of the curve from the diagonal is also assumed to reflect the difference between hypothetical, subjective "familiarity distributions" for the targets and lures. The farther the two distributions are from each other, the less likely it is that targets and lures will be confused. D', a measure of the distance between the familiarity distributions of targets and lures, was computed for each condition (Coombs, Dawes, and Tversky, 1970). Assuming that the underlying familiarity distributions are normal and of equal variance* five estimates of d' (for each condition) were obtained corresponding to the five hit-to-false alarm ratios that were free to vary. The five d's were averaged to obtain one d' measure for each condition, and the mean d's are shown in Table 2. From this table it can be seen that the recognition performance of subjects in the HH condition was greatest, and in the HU condition, the poorest. The fact that the difference between the d's for the HH and HU conditions is greater than the difference between the d's for the LL and LU conditions means that the addition of irrelevant low-frequency components during the test phase impairs recognition performance more than the addition of irrelevant high-frequency components. The results are not consistent with the perceptual masking hypothesis described earlier. The slightly greater interference produced by the addition of low-frequency components

---

*The assumptions are justified if the curves are symmetrical about the diagonal. The data of Fig. 15 appear to support the assumptions.

Table 2

D' Estimates for Experimental Conditions

Condition	D'
HU	.13
HH	.60
LU	.28
LL	.53

during the test phase may be interpreted as a bias toward the processing of the low-frequency components of complex scenes. The reason for the bias is not clear from the recognition data by themselves. However, if the bias were due to an observer's judgment about the relative importance of different types of physical information (e.g., coarse versus fine detail) in the scenes, then informativeness ratings for different areas of the pictures used in the recognition study might vary as a function of how the pictures are spatially filtered. Also, the bias might be reflected in the manner in which the spatial frequency content of local areas of the scenes controls an observer's scanning behavior. If so, the scanpaths of observers ought to vary as a function of the manner in which the scenes are spatially filtered. Judgments of informativeness and visual fixation patterns for a subset of the pictures used in the recognition study are described in the next two sections.

Informativeness Ratings for Local Areas of Complex, Real-life Scenes

In this study subjects ranked the informativeness of the local areas of pictures used in the recognition study. In order to determine how the rankings were related to the spatial frequency content of the local areas, the rankings were obtained under three different modes of presentation. The pictures were either unfiltered (U), low-pass filtered (L), or high-pass filtered (H), and the rankings obtained under each mode of presentation were compared (again, the filtering procedures described by Capell [1975] were used).

Method

Subjects. The subjects were 16 Miami University undergraduates enrolled in a perception course. They were given extra credit in the course for their participation in the experiment. They were divided into two groups, eight subjects to a group.

Materials. A subset (10) of the 20 slides used in the picture recognition experiment described in the previous section was used in the present study. Viewing conditions (viewing distance, spatial subtense of the projected image of the slides, etc.) were identical to those of the picture recognition study except that a second projector was used to superimpose the image of a grid transparency on each projected slide image. The grid of lines divided the stimulus slide into 24 equal areas (bins), 4 rows by 6 columns. The bins were numbered consecutively from 1-24 starting in the left, uppermost bin and proceeding across and down the grid.

Procedure. The subjects were tested in groups. When each

stimulus was presented, each subject judged the relative "informativeness" of each bin in the grid. The subject was given a sheet of paper with the ranks 1-24 listed across the top and instructed to place under each rank the number of the appropriate bin (1-24). Each subject began with the four or five most informative bins, then the four or five least informative bins, and finally the remaining bins. The subjects were not given a specific definition of "informativeness." The subjects were given approximately 4 min to rank the bins for each stimulus. When the last subject had completed the ranking task for a given stimulus, it was removed and the next one projected.

One group of subjects saw five (set A) of the 10 stimuli in the U presentation mode first and then the other five stimuli (set B) in the L presentation mode. The second group of subjects saw the five stimuli of set B in the U mode first and then the five stimuli from set A in the H mode. Both groups of subjects were given a 30-min break after the first five slides had been presented before they ranked the second set.

Results and Discussion

The correlations between the informativeness rankings for the H stimuli and their unfiltered counterparts are shown on the right in Table 3. The correlations were obtained in the following manner. For each filter version (H or U) of a picture, the average rank for each bin was computed by averaging the ranks of the eight subjects. The averages were then ranked from 1-24 and the ranks for the two filter conditions were correlated.

Table 3

Spearman Rank Order Correlation Coefficients

for Informativeness Ranks

Picture type	
Unfiltered with low-pass filtered	Unfiltered with high-pass filtered
Rho	Rho
0.21	0.59
0.57	0.76
0.88	0.67
0.92	0.72
0.14	0.85

Thus, the correlations on the right in Table 3 are between-group comparisons and represent the tendency for different subjects to rank the bins for H and U versions of a picture in the same way. The correlations between the informativeness rankings of H and U versions of the same picture are all relatively high.

The correlations between the informativeness rankings for the L stimuli and informativeness rankings for their U counterparts are shown on the left in Table 3. The correlations were computed in the same way as those between the H stimuli and their counterparts. The correlations are a measure of the tendency for different subjects to rank the informativeness of portions of a L or U version of a picture in the same way. Unlike the correlations between H and U versions of a picture, the correlations between L and U versions of a picture are not consistently high. Only two of the correlation coefficients (0.87 and 0.92) are as high as those for the H-U correlations.

The best explanation of the entire set of results is that subjects are biased in the direction of using high spatial frequency information when judging bin informativeness. Given this explanation, the reason why the H-U correlations are all consistently high is that both filter versions contain the same high frequency information. The low frequency information in the U scenes is simply ignored. At first glance, the set of L-U correlations does not seem to support the high frequency bias explanation. If the high frequency information is removed from a picture one would expect the ranks of bin informativeness to change, and L-U correlations should be low. However, this reasoning

is valid only if the low and high frequency content of bins within a picture is negatively correlated or not correlated at all. In the case of pictures with low L-U correlations (0.14 and 0.21) this second assumption appears to be justified because each picture contained only a few coarse features with fine details distributed widely throughout the picture. On the other hand, if the low and high spatial frequency content of bins within a picture is positively correlated, then the correlations between informativeness ranks for L and U stimuli would be high (artifactually) even with the high frequency bias. Due to their high spatial frequency bias subjects would use high frequency components to rank the bins of U scenes. But the ranks would not be different even if they were forced to rely on the low spatial frequency content of the bins (as they were when shown L stimuli) since the low and high frequency contents are positively correlated. Thus, it is possible to obtain a high L-U correlation for a picture even when subjects use the high spatial frequency content of the U picture and the low spatial frequency content of the L picture to rank bin informativeness, provided only that the low and high frequency contents of the bins are positively correlated. It is interesting that the pictures with high L-U correlations (0.88 and 0.92) had most (if not all) of their fine detail in regions where there were also coarse features (a horizon in one picture and a waterfall in the other).

To summarize, the high spatial frequency content of the local areas of visual scenes appears to be a better predictor of informativeness than the low spatial frequency content of the

areas. Given that informativeness judgments depend more on the high than on the low spatial frequency content of the regions of a scene, it seems paradoxical that recognition performance would be impaired more by the addition of low frequency "noise" during a recognition test than by the addition of high frequency "noise" (the picture recognition result described in the last section). Shouldn't performance be impaired more by information judged to be important or informative? Further discussion of this paradox is delayed until the next section after the visual scanpath data are described.

Visual Scanning of Complex, Real-life Scenes

Capell (1975) and Pantle (1975) described procedures which they used to obtain visual scanning data from human subjects as they viewed complex, real-life scenes. Briefly, scanning data were obtained from each of 12 subjects with the oculometer at Wright-Patterson Air Force Base. Each subject participated in a free-viewing task in which he simply inspected a real-life scene which was presented on a screen by a slide projector. During the 9 sec that the subject inspected each stimulus, electrical signals corresponding to his eye movements were recorded by the oculometer. Each subject saw fifteen pictures selected from the set of 20 slides used in the picture recognition study described in this report. Five stimuli were presented in the unfiltered (U) mode; five, in the low-pass (L) mode; and five, in the high-pass (H) mode. At the completion of the experiment, eye movement data in the form of analog electrical signals were low-pass filtered with a high-frequency cutoff of 4 Hz. The filtered signals were sampled at a 40-Hz rate to yield a total of approximately 360 samples of a subject's fixations per picture. At this point the data were prepared in two formats: (1) The digitized signals (or samples) were amplified and fed into an X-Y plotter to reconstruct the scanpaths of the subject. The scale of the X-Y plots was adjusted to match the size of 5 x 7 in prints of the stimulus scenes which the subject had viewed. The scanpaths were used as overlays on the prints to determine directly the points fixated in each scene. It was also possible to get some appreciation of the areas of each scene fixated most often by visual inspection

of the scanpath data. Fixations tended to cluster in some areas more than others. (2) In order to obtain a more precise quantitative estimate of the percentage of time that a subject fixated different areas of each stimulus scene, each scene was divided into 24 equal areas (bins) as in the informativeness study. Using calibrated values, the position of each sample fixation was classified as being in one of the 24 bins. The number and percentage of fixations in each bin was then readily determined.

Statistical analyses of fixation data. A subset of the scanning data collected by Capell (1975) and Pantle (1975) was chosen for further analyses here according to two criteria. (1) Scanning data were selected if informativeness data were available for the same pictures. (2) The fixation data were reliable.

The first selection criterion was used in order to make it possible to correlate eye fixations with informativeness rankings. The second criterion was used in order to obtain the best possible data for further analyses. The accuracy and reliability of the two data records (bin counts and scanpaths) were checked in several ways after it became apparent that the two data records were discrepant; i.e., the bin counts did not match well with the number and/or location of fixations on the scan-paths. Either one or the other or both of the records had to be unreliable. Efforts were made to determine the source of the unreliability.

The bin counts for each subject and picture [those chosen by

criterion (1) above] were compared by visual inspection to prints of each picture to determine the reasonableness of the bin counts. The data of one subject were discarded on the grounds that the majority of his fixations were in the middle of the picture irrespective of the scene being viewed. There were three other cases in which a subject's data for a single H or L picture were eliminated because the picture was apparently shown upside down. On the whole, the bin count data seemed reasonable. While the data sets for some pictures were far in excess of or short of the expected 360 samples per picture, each of the data sets chosen for further analyses contained approximately 360 samples.

The X-Y plots were next checked for their accuracy and reliability. The plots were first checked for any regular overall shift in the fixation positions that could have been introduced due to a change in calibration at the time of recording or at the time of plotting. Such a change in calibration would have resulted in a consistent tendency for the plotted points to move generally nearer one edge of the plot with each successive plot for a given subject. Appreciable shifts were observed along the vertical axis in the data for two subjects, as shown by the fact that there was a tendency for points to collect at the top of the picture. When the X-Y plots were checked against prints of the pictures, oftentimes it appeared that the subjects' fixations were reasonable only if the plots were shifted relative to the prints. Because the bin counts were fairly reliable and the scanpaths were not, it was assumed that the problem with matching scanpaths to pictures was due to errors at the time of plotting.

Apparently, the calibration of the plotter drifted when the scanpaths were produced. The X-Y plots were not used as a criterion for selecting reliable data and only the bin count data (with the deletions described above) were further analyzed.

For each picture viewed in the same presentation mode (either U, L, or H), the percentage of fixations for each bin was averaged across subjects. The averages were then used to determine the correlations between the fixations of U stimuli and their filtered counterparts. The correlations are reported in Table 4. Both Spearman Rank Order and Pearson Product Moment correlations are reported as there were some differences between them. It should be noted that the correlations in Table 4 are quite high despite the fact that they represent the agreement among the fixation behaviors of <u>different</u> subjects viewing <u>different</u> filter versions of the same pictures. The Spearman L-U correlations (average Spearman Rho = 0.85) are higher than the Spearman H-U correlations (average Spearman Rho = 0.68). The Pearson L-U correlations (average Pearson $\underline{r}$ = 0.71) are also higher than the Pearson H-U correlations (average Pearson $\underline{r}$ = 0.64), although the difference is not as great as that between the corresponding Spearman coefficients.

Unlike the informativeness rankings, there is no evidence that high spatial frequency information guides scanpaths more than low spatial frequency information. If anything, the results suggest that low spatial frequencies are a better predictor of eye fixations.

Table 4

Correlations Between Filtered and Unfiltered Versions of Pictures: Eye Fixation Data

Picture type			
Unfiltered with low-pass filtered		Unfiltered with high-pass filtered	
Spearman	Pearson	Spearman	Pearson
0.88	0.74	0.73	0.86
0.81	0.56	0.58	0.41
0.82	0.71	0.72	0.80
0.72	0.73	0.60	0.52
0.83	0.79	0.75	0.63

AD-A049 127 MIAMI UNIV OXFORD OHIO DEPT OF PSYCHOLOGY F/G 6/4

RESEARCH ON VISUAL PERCEPTION OF COMPLEX AND DYNAMIC IMAGERY.(U)

NOV 77 A J PANTLE F33615-76-C-5006

UNCLASSIFIED AMRL-TR-77-83 NL

2 of 2

ADA049127

END

DATE FILMED

3-78

DDC

Fixation and Informativeness Data Compared

In order to address more directly the question of how eye fixations are related to "informativeness," the fixation data were compared to the informativeness rankings. For each picture viewed and ranked, a Spearman Rank Order correlation between eye fixations (percent bin count averaged across subjects) and averaged ranked informativeness was computed. The correlations are shown in Table 5. The highest average correlation between fixation time and informativeness obtains for L pictures, next highest for H pictures, and finally for U pictures. In other words, it appears that when information is removed from a picture there is more agreement between what is looked at and what is considered as informative.

If there were a tendency for low spatial frequencies to control eye fixations and a tendency for high spatial frequencies to determine informativeness (as suggested in earlier sections), then the lowest correlations between informativeness and fixation time should be obtained for U pictures because both sets of frequencies are present in U pictures. With either H or L pictures, the single set of spatial frequencies present in a picture would have to determine both the informativeness of bins and the fixation time for the bins, and the correlations between informativeness and fixation time should be higher than for U pictures. If, on the other hand, both fixation time and informativeness depended upon the same set of spatial frequencies (an assumption not supported by the results in earlier sections), then the correlations between fixation time and informativeness

Table 5

Spearman Rank Order Correlations Between Fixation Bin Counts and Informativeness Ranks

	Picture type		
	Unfiltered	Low-pass filtered	High-pass filtered
	0.02	0.41	0.52
	0.59	0.41	0.47
	0.73	0.77	0.35
	0.24	0.40	0.46
	0.19	0.54	0.41
	0.42		
	0.06		
	0.46		
	0.64		
	0.36		
Average	0.37	0.51	0.41

would be as high for the U pictures as for the set of filtered pictures that retained the spatial frequencies determining fixation time and informativeness.

As can be seen in Table 5, the correlations for the U pictures are generally the lowest of the three sets. Moreover, half of the correlations for the U pictures are not statistically significant whereas only one of the other correlations is not statistically significant. Therefore, it can be tentatively concluded that fixation time and informativeness depend upon different spatial frequency components. Thus, the direct comparison of fixation and informativeness data is consistent with the conclusions reached in the separate analyses of the fixation and informativeness data.

Conclusions

The conclusions summarized in this section should be regarded as suggestive because they are based to a large extent on correlational data and the differences among correlation coefficients were oftentimes small. Also, the conclusions might reasonably be expected to hold only for stimulus scenes of the type studied. Nonetheless, so little is known about the physical variables which control the manner in which information in complex scenes is processed that the tentative conclusions offered here may serve as valuable starting points for future research.

The informativeness study suggests that the high spatial frequency content of the local areas of real-life scenes is a better predictor of the judged informativeness of the areas than is the low frequency content. The scanning data suggest that the low spatial frequency content of the local areas of real-life scenes is a better predictor of the percentage of time that a local area is fixated than is the high spatial frequency content. In the picture recognition experiment it was found that the addition of irrelevant low spatial frequency content ("noise") to stimuli during the test phase impaired recognition performance more than the addition of irrelevant high spatial frequency content.

The results of the picture recognition and eye fixation studies appear to fit together. Low spatial frequency information may have impaired recognition performance more because the eyes (and attention?) of subjects were drawn to low frequency "blotches" in scenes during the test phase, and consequently proper weight was not given by the subjects to other high spatial frequency

areas which they had inspected carefully and could have used for retrieval cues during the study phase.

Superficially at least, the relationship between the results of the informativeness study on the one hand and the results of the picture recognition and eye fixation studies on the other is puzzling. It seems to make little sense that eye fixations should have a greater tendency to be controlled by spatial components of scenes (low) which subjects judge to be less important (informative) than other components (high). Moreover, it seems to make little sense that frequency components (low) which subjects judge to be less important (informative) than other components (high) should have a greater tendency to interfere with recognition performance when they are added as "noise" during tests for recognition. The paradoxes may stem from the nature of the informativeness task itself, and the particular definition of informativeness which the subjects adopted for themselves. In retrospect it seems logical that subjects would base their ranks on the amount of fine visual detail (high spatial frequency energy) in the bins of each scene because they were asked to rank the relative informativeness of bins <u>within the same picture</u>. Perhaps the low spatial frequency content of bins (low frequency "blotches") would have influenced the informativeness ranks more had the subjects been asked to rank the informativeness of a bin on the basis of how much it would permit one to discriminate one scene from another or how important it would be in a later recognition test with many different pictures. Given one of the alternative operational definitions of informativeness,

informative bins might become noninformative and vice versa and the paradox of why unimportant low frequency "noise" impairs recognition performance more than important high frequency "noise" and the paradox of why the eyes are drawn to unimportant or noninformative parts of scenes might disappear.

Research on Velocity Coding

At present human psychophysical data are not adequate for deciding between various alternative mechanisms for velocity coding. Part of the inadequacy stems from the fact that past studies have used stimuli whose spatial frequency content was unknown. Also, past experiments simply were not designed to test alternative velocity coding schemes.

Hypothetical Principles of Velocity Coding

Perhaps the simplest code for velocity would be one in which the output of single movement-sensitive cells increases as stimulus velocity increased (additive model). In other words, high impulse rates would be associated with fast-moving stimuli; low rates, with slow-moving stimuli. In humans, subjective speed would be determined by the absolute level of the output of movement-sensitive cells. Movement-sensitive cells which possess the response characteristics of the type required by the additive model have been found in the visual systems of some species. For example, Grusser et al. (1968) determined the relationship between the angular velocity of a stimulus and the average discharge rate of movement-sensitive cells. Nonperiodic stimuli were moved through the excitatory receptive field of each neuron and the impulse rate was measured during the traversal of the receptive field. The average discharge frequency of the neurons increased with the angular velocity of the stimulus. The relationship between impulse frequency and angular velocity was described by a power function and it held over a range of 2-3 log units. The wide velocity range over which response rate increases

would seem to provide an adequate physiological mechanism for additive coding of velocity.

While the additive model is attractive because of its simplicity, and while it may describe velocity coding in some lower species, there are a number of experimental results which cast doubt on its validity for higher vertebrate systems. Unlike the movement-sensitive neurons of lower species, directionally selective movement cells in vertebrates (Barlow, Hill, and Levick, 1964; Hubel and Wiesel, 1965; Ganz and Lange, 1973) respond to a narrow range of velocities. The responses are greatest at an intermediate velocity and less at both slower and faster velocities. Moreover, the optimal velocity for stimulating a cell varies from one cell to the next. Hubel and Wiesel (1965) report that the most effective rates of movement for complex cells of the cat varied from 0.1°/sec up to about 20°/sec. As a consequence of their velocity tuning, the distribution of cells affected by a stimulus moving at one speed will be different than the distribution of cells affected by a stimulus moving at another speed. For example, a stimulus moving at 20°/sec would maximally excite cells tuned to a stimulus velocity of 20°/sec and cause little or no activity in cells tuned to a velocity of 1°/sec. The reverse would be true for a stimulus moving at 1°/sec. As stimulus speed is changed, the responses of some cells <u>substitute</u> for those of other cells. Therefore, it is possible that perceived velocity depends upon the relative activities among velocity-tuned cells, much like perceived color depends upon the relative excitation of different cone systems. The

substitutive model of velocity perception is also supported by a human psychophysical experiment of Pantle and Sekuler (1968). They demonstrated the existence of velocity-tuned mechanisms in human vision by means of a detection experiment in which a selective adaptation procedure was employed.

Effect of Motion Adaptation on the Perceived Speed of Sinusoidal Gratings

Even though physiological studies and the psychophysical experiment of Pantle and Sekuler (1968) reveal the existence of velocity-tuned mechanisms in higher vertebrate systems, they do not directly demonstrate that velocity judgments and discriminations are governed by the substitutive principle. More direct tests of the additive and substitutive models might be provided by psychophysical experiments in which a subject first adapts to a moving grating and afterwards judges the speed of a moving test grating, sometimes moving slightly faster than the adapting grating and sometimes moving slightly slower. If it is assumed that the response of a motion-sensitive mechanism diminishes with prolonged stimulation (see Barlow and Hill, 1963), then the outcome of the proposed experiment can be predicted. According to the additive model, the general loss of responsiveness of motion-sensitive cells that is brought about by adaptation to a moving grating ought to decrease the absolute level of activity that is produced by a test grating moving at any speed. Hence, the apparent speed of all test gratings, regardless of their actual speed, should decline following adaptation to a moving grating. The substitutive model makes a different prediction. A test grating moving at a particular speed would be expected to produce activity in a distribution of velocity-tuned mechanisms. Of the distribution of mechanisms responding to the test grating those tuned to slower velocities ought to be adapted most by an adapting stimulus moving slower than the test grating. Consequently, adaptation to a slow-moving grating ought to increase

the apparent speed of a test grating moving at a slightly faster speed. Similarly, adaptation to a fast-moving grating ought to make a slower test grating appear slower than it really is. Measurements of the perceived velocity of test gratings can be made with a matching technique. The physical speed of a comparison grating in an unadapted area of the retina is adjusted until it matches the physical speed of a test grating in an adapted area of the retina.

Method

The subject's task was to adjust the speed of one sinusoidal grating (comparison grating) to match the speed of a second sinusoidal grating (test grating) drifting in the opposite direction. The matching speed of the comparison grating provided a measure of the perceived speed of the test grating. Velocity matches were made after the subject had adapted to gratings (adapting gratings) moving at various speeds and located in the same area of the visual field as the test grating in order to determine the effects of velocity adaptation on the perceived speed of the test grating.

The comparison and test gratings were generated on the faces of two oscilloscopes (Tektronix 5000 series scopes with P-31 phosphors), one positioned above the other. During a test period the test grating was displayed on the upper scope and the comparison grating was displayed on the lower scope, such that the test and comparison gratings were imaged on the inferior and superior retinal areas, respectively, of the subject's eyes. During an adapting period the adapting grating was displayed on

the upper oscilloscope, and the lower scope was uniformly illuminated.

The technique described by Enroth-Cugell and Robson (1966) was used to generate the gratings. The speed of the adapting, test, and comparison gratings was the only characteristic which was changed during the course of the experiment. The gratings were equal on other dimensions. The spatial frequency of the gratings was 1 c/deg. They subtended 2° 52' vertically and 3° 50' horizontally. The contrast of the gratings was 0.26, and their space-average luminance was 5.0 mL. The luminance of the uniform field on the lower scope during adaptation periods was also 5.0 mL. The gratings were viewed through rectangular holes in a cardboard surround, whose color approximately matched the scope faces and whose luminance was 3.8 mL. The distance between the lower edge of the upper grating and the upper edge of the lower grating was 2° 9'. The subject fixated a point midway between the two edges.

An adaptation-test paradigm was used in the experiment. The subject inspected an adapting grating continuously during the initial 3 mins of each experimental session, and, thereafter, during intervals between successive test periods. The adapting intervals were 10 secs long and they alternated with the 2.0-sec test periods. The changeover from the adapting display to the patterns present during the test period required 3.5 secs. During the first 0.25 sec, the contrast of the adapting grating decreased linearly from 0.26 to 0.0. During the last 0.25 secs of the changeover interval the contrast of the comparison and

test gratings increased linearly from 0.0 to 0.26. During the middle 3 secs, the contrast of all patterns was 0.0 (i.e., spatially uniform).

The adapting grating moved at one of three different drift rates--3.7, 4.8, or 5.6 Hz--to produce three different conditions of velocity adaptation. The adapting gratings drifted in the same direction as the test gratings. In addition, a fourth adapting condition (control condition) was used in which the adapting grating was replaced by a spatially uniform field (5.0 mL).

For each of the four adapting conditions, velocity matches between the test and comparison gratings were obtained, using a tracking procedure. After each test period the subject (using a servo-device) increased or decreased the speed of the comparison grating depending upon whether it had appeared to move slower or faster than the test grating. For the initial test period, the speed of the comparison grating was set well above or below that of the test grating. The physical speed of the test grating was always 4.8 Hz, although, as the results show, its apparent speed depended upon the subject's condition of adaptation. A tracking series was terminated after eight reversals of the direction in which the subject adjusted the speed of the comparison grating. The mean of the last six reversals provided an estimate of the perceived speed of the test grating. Ten estimates of the perceived speed of the test grating were obtained in each adapting condition for each of two subjects (LMP and PLS); six estimates for each of two other subjects (PAJ and AJP). One estimate of perceived speed of the test grating was obtained for each

adapting condition during each experimental session. At least 4 mins elapsed between runs in the different adapting conditions. The order of presentation of the four adapting conditions was counterbalanced for each subject across experimental sessions.

Results and Discussion

The results of the experiment are summarized in Table 6. Each of the first four panels in the table is the results for an individual subject. Each of the first four panels gives the mean and standard deviation of the individual estimates of perceived test grating speed in one of the four adapting conditions. The last panel in the table gives the group means for each adapting condition and the standard deviation of the means of individual subjects about each of the respective group means.

In the control condition, the speeds of the comparison and test gratings appeared equal when their physical speeds were approximately equal. The overall average matching speed of 4.9 Hz is, within the limits of experimental error, equal to the physical speed of the test grating. However, in all other adapting conditions (conditions of velocity adaptation), an apparent speed match was not achieved when the physical speeds of the comparison and test gratings were equal. The main effect of adapting condition on matching speeds was found to be significant in a repeated measures analysis of variance [$\underline{F}(3, 9) = 23.33$, $\underline{p} < 0.001$]. In general, the physical speed of the comparison grating was slower than that of the test grating when their perceived speeds matched. This means that the inspection of a moving adapting grating reduced the apparent

Table 6

Mean and Standard Deviation of Estimates of Perceived Test Grating Speed in Different Adapting Conditions

Subject		Adapting condition			
		Control	3.7 Hz	4.8 Hz	5.6 Hz
LMP	$\bar{X}$ (N = 10)	4.75	3.81	3.42	3.18
	S.D.	0.33	0.30	0.21	0.18
PLS	$\bar{X}$ (N = 10)	4.92	4.23	4.24	4.20
	S.D.	0.29	0.33	0.45	0.39
PAJ	$\bar{X}$ (N = 6)	5.18	4.45	3.74	3.22
	S.D.	0.36	0.28	0.34	0.26
AJP	$\bar{X}$ (N = 6)	4.75	4.13	3.49	3.18
	S.D.	0.20	0.54	0.24	0.30
Overall mean		4.90	4.16	3.72	3.45
Overall S.D.		0.20	0.27	0.37	0.50

speed of the test grating. The reduction occurred irrespective of adapting speed. This finding is consistent with the additive principle of velocity coding, but not with the substitutive principle. Under the substitutive principle there should have been no reduction in matching speed in the 4.8-Hz adapting condition and an increase of matching speed in the 3.7-Hz adapting condition.

In general, the magnitude of the reduction of the perceived speed of the test grating after adaptation to a moving grating was not the same for all adapting speeds. The overall, average matching speed was highest for the 3.7-Hz adapting condition, intermediate for the 4.8-Hz condition, and lowest for the 5.6-Hz adapting condition. This trend is present in the data of all subjects except PLS. The relative order of the reductions in the perceived speed of the test grating is consistent with either the additive or the substitutive principle of velocity coding. However, as described above, other aspects of the data are not consistent with the substitutive principle. Under the additive principle, one would expect the reductions in the perceived speed of the test grating to be directly related to the adapting speed because the output, and consequently the adaptation of movement-sensitive cells, is assumed to be greater for faster speeds. In summary, the additive principle of velocity coding provides a better overall account of the velocity adaptation results than does the substitutive principle.

Spatial Frequency and Perceived Velocity

In the previous velocity coding experiment the spatial frequencies of the sinusoidal gratings whose speeds were matched were the same. In the present experiment subjects compared the speeds of sinusoidal gratings of different spatial frequencies. The purpose of the experiment was to evaluate the influence of the spatial frequency of moving stimuli on their perceived speed.

Method

Subjects. The subjects were two trained psychophysical observers who were paid for their participation. Both subjects were naive with respect to the experimental goals.

Procedure. The stimulus display was identical to that used in the previous velocity coding experiment except that the test grating was a 3-c/deg sinusoidal grating. The velocity (the visual angle through which a point on the stimulus moves in a second) of the test grating was 2.5 deg/sec which corresponds to a temporal drift rate (the number of periods of the pattern moving by a fixed point in a second) of 7.5 Hz. The Method of Constant Stimuli was used to determine that speed of the comparison grating (1 c/deg) which appeared equal to that of the test grating. The test and comparison gratings were presented simultaneously (a trial) every 4-8 sec. They were turned on for 2.5 sec inclusive of a 250-msec period during which the contrast rose linearly from zero to the maximum and a 250-msec period during which the contrast fell linearly from the maximum to zero. The comparison grating was moved at one of 10 different speeds on each trial, and after each trial the subject indicated whether he

thought the comparison grating moved faster or slower than the test grating. The 10 speeds of the comparison grating were 1.0, 1.8, 2.5, 3.5, 4.5, 5.5, 6.5, 7.5, 7.8, and 8.0 deg/sec. The corresponding temporal drift rates are numerically equal to the speeds. Each of the comparison speeds was presented 32 times for a total of 320 trials per subject. Each subject participated in eight sessions of 40 trials each. Within each session each comparison speed was presented four times, and all ten comparison speeds were presented _n_ times before any speed was presented for the _n + 1_ th time. The order of presentation of the 10 speeds within any 10-trial block was random.

Results

For each comparison speed, the number of "faster" responses was tabulated and converted to a percent. Because the results of both observers were essentially the same, the mean percentage of "faster" responses was computed for each comparison speed. The means are shown in Fig. 16. The figure shows that the percentage of "faster" responses grew monotonically with the drift rate of the comparison grating. It can also be seen from the figure that the comparison grating was judged to be moving faster than the test grating on approximately half the trials and slower on the other half when its temporal drift rate was 4.5 Hz. This value is the point of subjective equality (PSE) or matching drift rate. Converted to a speed measure, the PSE was 4.5 deg/sec.

The dashed line labeled "temporal frequency match" in Fig. 16 corresponds to the PSE which would have been obtained if the

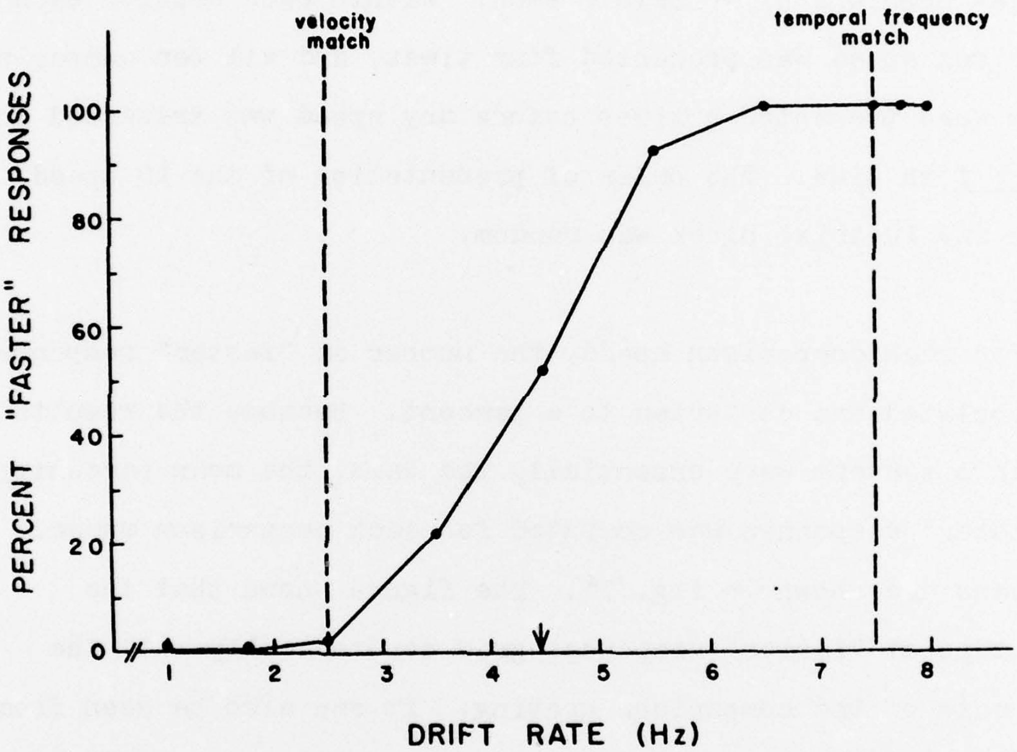


Fig. 16. Percentage of faster responses as a function of the drift rate of the comparison grating. The arrow represents the actual point of subjective equality (PSE), and the dashed lines correspond to the PSE's based on a velocity match and a temporal frequency match, respectively.

perceived velocities of the test and comparison gratings were equal when their physical temporal frequencies (drift rates) were equal. Clearly, temporal drift rate alone was not the basis or criterion for the subjects' judgments of perceived velocity since the comparison grating was always judged to be moving faster than the test grating when their temporal drift rates were physically equal. The dashed line labeled "velocity match" in Fig. 16 corresponds to the PSE which would have been obtained if the perceived velocities of the test and comparison gratings were equal when their physical velocities were equal. Clearly, physical velocity alone was not the basis or criterion for the subjects' judgments of perceived velocity since the comparison grating was judged almost without exception to be moving slower than the test grating when their physical velocities were equal. The actual PSE is some compromise between or combination of the PSE based on physical velocity and the PSE based on temporal frequency (or drift rate).

The Perceived Velocity of Complex Gratings

In the control (no-adaptation) condition of the first velocity coding experiment it was found that the perceived speeds of two sinusoidal gratings of the same spatial frequency matched when their physical velocities were equal. In the second velocity coding experiment it was found that the perceived speed of an f grating matched that of a 3f grating when the physical speed of the f grating was approximately 80% faster than that of the 3f grating; i.e., the subjective match was based on some compromise between a physical velocity match and a physical temporal frequency match. It is not clear what would happen if subjects matched the perceived speed of an f grating with that of a complex, f + 3f grating. If the subjects simply ignored the 3f component, then the perceived speeds of the gratings would presumably match when their physical speeds were equal. If the subjects ignored the f component of the complex grating, then the requirements for a subjective match would presumably be the same as for the case when a subject matches the perceived speed of an f grating to that of a 3f grating, i.e., the compromise described above. If both components of the complex grating determine its perceived speed, then the physical f-grating speed required for a subjective match ought to be somewhere between the speeds required for subjective matches obtained with the f and 3f components alone.

Method

The subjects were two trained psychophysical observers, PAJ and CLF. Subject CLF was naive with respect to experimental

goals, while subject PAJ was not. The stimulus display was identical to that employed in the previous velocity coding experiments except that there were seven different test gratings, two of which were sinusoidal gratings. One sinusoidal grating had a spatial frequency of 1 c/deg (f) and the other, 3 c/deg (3f). The contrast of the sinusoidal gratings was 0.27. There were five complex test gratings, each with a 1-c/deg fundamental and a 3-c/deg harmonic (f + 3f gratings). The amplitude ratio (f/3f) of the components of the complex gratings was different for each of the five complex gratings: (a) 0.27/0.09, a 3:1 ratio; (b) 0.27/0.18, a 3:2 ratio; (c) 0.27/0.27, a 1:1 ratio; (d) 0.18/0.27, a 2:3 ratio; and (e) 0.09/0.27, a 1:3 ratio. The test gratings were moved at one of three different velocities-- 1.0, 2.5, and 4.0 deg/sec. At these velocities the corresponding temporal frequencies (drift rates) of the f grating or the f component of the complex grating were 1.0, 2.5, and 4.0 Hz; the corresponding temporal frequencies of the 3f grating or the 3f component of the complex grating were 3.0, 7.5, and 12.0 Hz. Five estimates of the speed (PSE) of the comparison grating (1 c/deg) which was required to match the speed of the test grating were obtained for each of the seven test gratings at each of the three test speeds. The test and comparison gratings were presented simultaneously as in the second velocity coding experiment, and the tracking procedure used in the first velocity coding experiment was used to obtain the PSE estimates. A total of 105 PSE estimates (five replications of the 21 test conditions) was collected for each subject. For each replication

the order in which the 21 conditions were run was random.

Results and Discussion

The results of the experiment are shown in Figs. 17 and 18 for subjects CLF and PAJ respectively. For each subject mean PSE's (in Hz) are plotted for each of the 21 test conditions. The test conditions in which different test gratings were employed are represented along the abscissa, with the two sinusoidal grating conditions at the extreme right and left. Each of the three sections of each figure contains data for a different test grating speed: the top section, data for a 4.0-deg/sec test speed; the middle section, data for a 2.5-deg/sec test speed; and the bottom section, data for a 1.0-deg/sec test speed.

Consider first the main effect of test grating speed. It is straightforward. It can be seen from the figures that the PSE increased at each amplitude ratio as the test velocity was increased from 1.0 to 2.5 deg/sec. Likewise, the PSE's increased when the test velocity was increased from 2.5 to 4.0 deg/sec. The main effect of test grating velocity on PSE was statistically significant [$\underline{F}(2, 2) = 1194.67$, $\underline{p} < 0.001$], and the relationship simply means that the perceived speed of the test grating is positively related to its physical speed.

The main effect of amplitude ratio was also significant [$\underline{F}(6, 6) = 8.98$, $\underline{p} < 0.01$]. Even though the velocity of the test gratings is constant for the seven conditions within each section of the figures, the PSE's change. In terms of absolute size, the differences among the PSE's may seem quite small, but

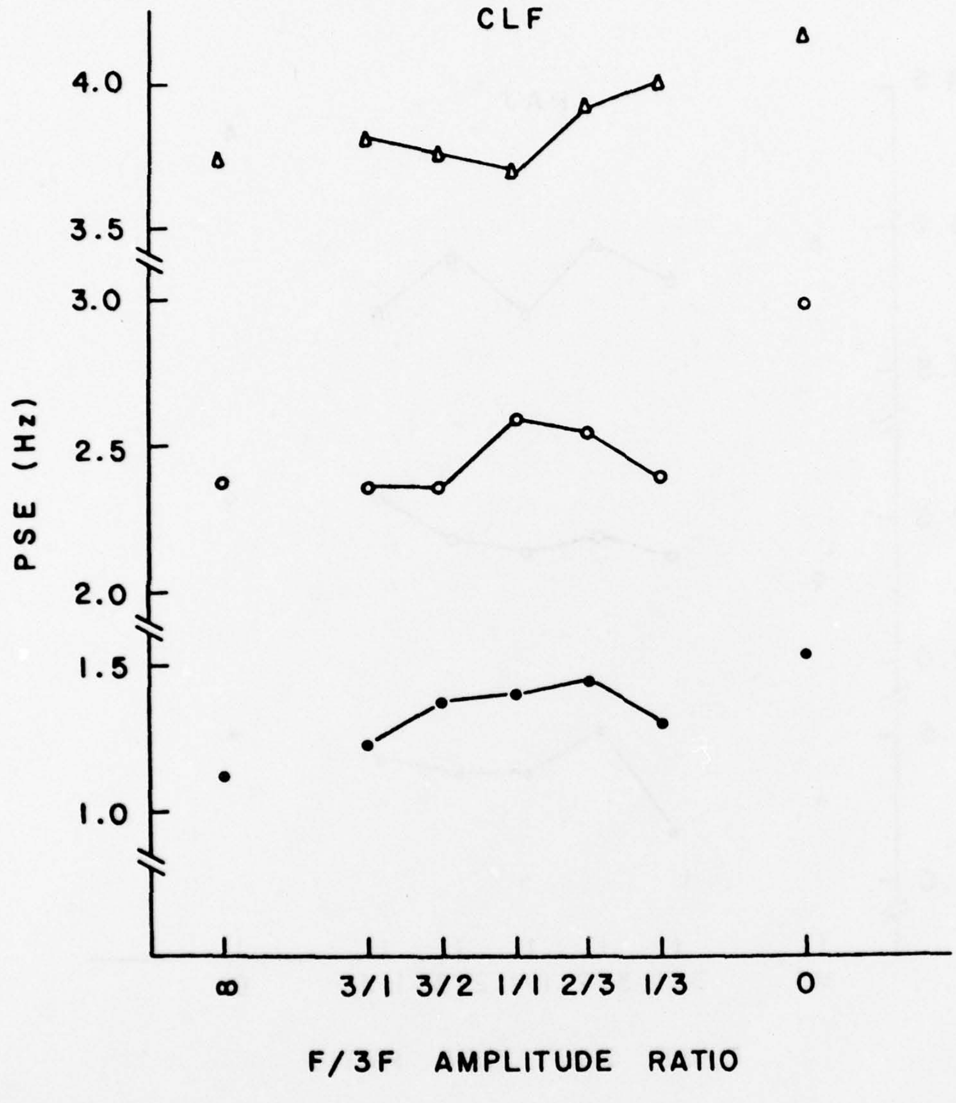


Fig. 17. Point of subjective equality (PSE) as a function of the relative amplitudes of the spatial frequency components of test gratings and of the physical speed of the test gratings. Physical test velocity of 1 deg/sec: ●'s; 2.5 deg/sec: o's; 4 deg/sec: Δ's. Subject CLF.

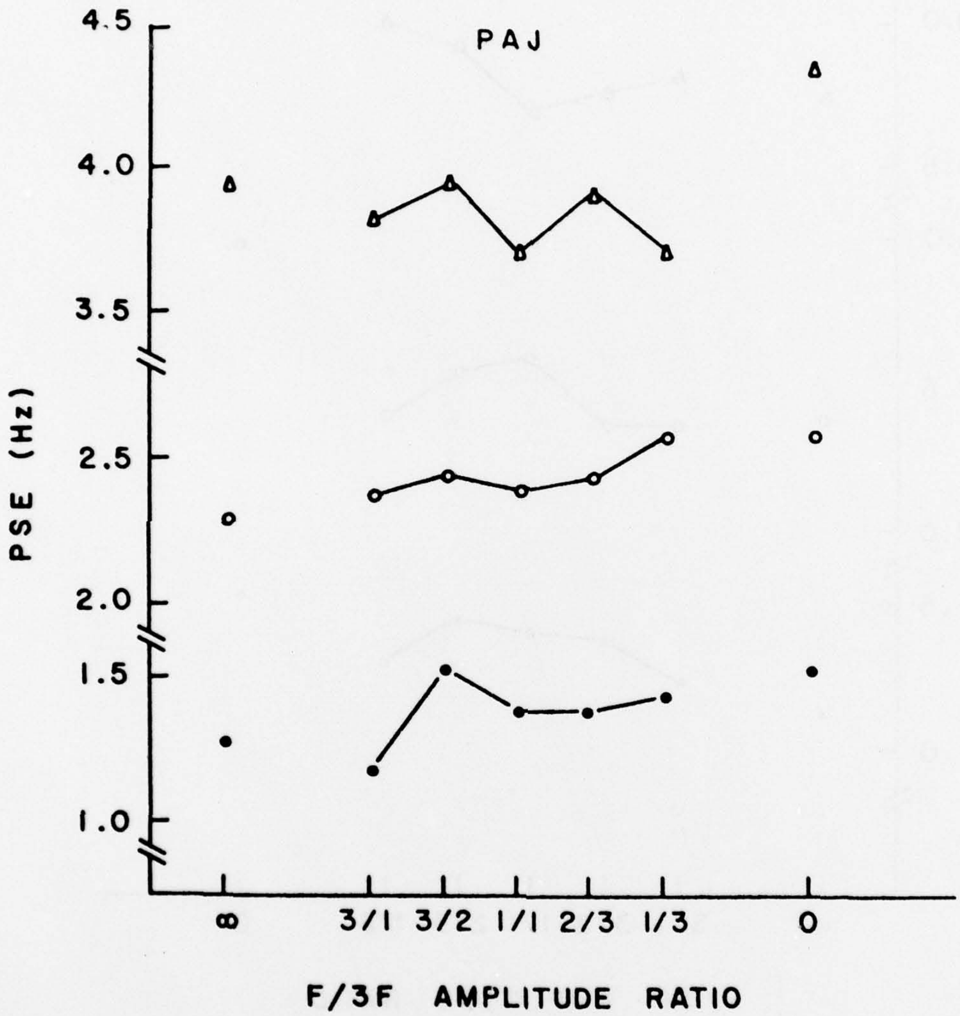


Fig. 18. Point of subjective equality (PSE) as a function of the relative amplitudes of the spatial frequency components of test gratings and of the physical speed of the test gratings. Physical test velocity of 1 deg/sec: ●'s; 2.5 deg/sec: o's; 4 deg/sec: Δ's. Subject PAJ.

the differences are statistically significant and it must be remembered that difference thresholds for the velocity of stimuli like those used in the present experiment are of the order of 1-2% (Pantle, 1975). The PSE's for the 3f test grating were consistently higher than those for the f test grating, indicating that the 3f test grating appeared to move faster than the f grating. This result is a replication of the findings in the second velocity coding experiment. Again, each of the subjective speed matches was based on a compromise between a physical velocity match and a temporal frequency match. With the exception of the PSE's obtained with the 4.0-deg/sec test speed for subject PAJ, the PSE's for the complex gratings are generally higher than those for the f test grating and lower than those for the 3f test grating. This result means that the addition of the third harmonic to the f test grating makes it appear to move faster, but not as fast as a 3f grating moving at the same physical speed. As described in the introduction to this experiment, the fact that the PSE's for the complex gratings are intermediate between those for the f and 3f test gratings is evidence that both components of the complex grating determine its perceived speed.

The increase of the PSE's caused by the addition of a third harmonic to the f test grating did not appear to vary systematically with the relative contrast of the third harmonic; i.e., the functions in Figs. 17 and 18 are on the average relatively flat for amplitude ratios from 3:1 through 1:3. This result suggests that the contribution made by the temporal

frequency of the third harmonic to the perceived speed of the complex grating is relatively independent of contrast as long as the contrast is suprathreshold.

Pantle (1975) has shown how the motion aftereffect phenomenon can be used to study the response properties of the transient visual channel. The magnitude of the motion aftereffect, and by inference the response of the transient channel, has been found to be relatively independent of stimulus contrast (Keck, Palella, and Pantle, 1976) and to be controlled by the temporal frequency of moving sinusoidal gratings rather than by their velocity per se (Pantle, 1974). The transient visual channel has response properties which make it an ideal candidate to serve as the basis for the contribution of the temporal frequency of the third harmonic to the perceived speed of the complex grating in the present experiment. More generally, the tendency of subjects to use a temporal frequency match as a criterion for matching the perceived speeds of moving gratings may reflect the operation of the transient visual channel. If so, the tendency of subjects to use a physical velocity match as a criterion for matching the perceived speeds of moving gratings must reflect the operation of different mechanisms. Whatever the nature of those mechanisms, their response would presumably be controlled by the velocity of moving sinusoidal gratings rather than by their temporal frequency.

References

Anstis, S. Phi movement as a subtraction process. Vision Research, 1970, 10, 1411-1430.

Barlow, H. B., Fitzhugh, R., and Kuffler, S. W. Change of organization in the receptive fields of the cat's retina during dark adaptation. Journal of Physiology (Lond.), 1957, 137, 338-354.

Barlow, H. B. and Hill, R. M. Evidence for a physiological explanation of the Waterfall phenomenon and figural after-effects. Nature (Lond.), 1963, 200, 1345-1347.

Barlow, H. B., Hill, R. M., and Levick, W. R. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. Journal of Physiology (Lond.), 1964, 173, 377-407.

Bell, H. H. and Lappin, J. S. Sufficient conditions for the discrimination of motion. Perception & Psychophysics, 1973, 14(1), 45-50.

Braddick, O. The masking of apparent movement in random-dot patterns. Vision Research, 1973, 13, 355-369.

Braddick, O. A short-range process in apparent motion. Vision Research, 1974, 14, 519-527.

Breitmeyer, B. G. A relationship between the detection of size, rate, orientation and direction in the human visual system. Vision Research, 1973, 13, 41-58.

Breitmeyer, B. G. and Ganz, L. Implications of sustained and transient channels for theories pattern masking, saccadic suppression, and information processing. Psychological Review, 1976, 83(1), 1-36.

Capell, C. M. Effects of alteration of spatial frequency content of complex scenes on human visual scan patterns. Unpublished Master's Thesis, Air Force Institute of Technology, WPAFB, 1975.

Coombs, C. H., Dawes, R. M., and Tversky, A. Mathematical psychology: An elementary introduction. Englewood Cliffs, New Jersey: Prentice-Hall, 1970.

Enroth-Cugell, C. and Robson, J. G. The contrast sensitivity of retinal ganglion cells of the cat. Journal of Physiology (Lond.), 1966, 187, 517-552.

Ganz, L. and Lange, A. Changes in motion sensitivity of cat visual cortex neurons during the course of dark adaptation. Presented at the Spring Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida, 1973.

Green, D. M. and Swets, J. A. Signal detection theory and psychophysics. New York: John Wiley & Sons, 1966.

Grusser, O.-J., Grusser-Cornehls, U., and Licker, M. D. Further studies on the velocity function of movement detecting class-2 neurons in the frog retina. Vision Research, 1968, 8, 1173-1185.

Harmon, L. and Julesz, B. Masking in visual recognition: Effects of two-dimensional filtered noise. Science, 1973, 180, 1194-1197.

Hubel, D. H. and Wiesel, T. N. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. Journal of Neurophysiology, 1965, 28, 229-289.

Julesz, B. _Foundations of cyclopean perception_. Chicago: University of Chicago Press, 1971.

Keck, M., Palella, T., and Pantle, A. Motion aftereffect as a function of the contrast of sinusoidal gratings. _Vision Research_, 1976, _16_, 187-191.

Kulikowski, J. J. and Tolhurst, D. J. Psychophysical evidence for sustained and transient detectors in human vision. _Journal of Physiology_, 1973, _232_, 149-162.

Loftus, G. R. and Bell, S. M. Two types of information in picture memory. _Journal of Experimental Psychology: Human Learning and Memory_, 1975, _104_, 103-113.

Mitchell, D. E., Freeman, R., Millodot, M., and Haegerstrom, G. Meridional amblyopia: Evidence for modification of the human visual system by early visual experience. _Vision Research_, 1973, _13_, 535-558.

Pantle, A. J. _Visual information processing of complex imagery_. AMRL-TR-74-43. Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratory, July 1974 (AD 785 366).

Pantle, A. J. _Research on the recognition and analysis of complex and dynamic imagery_. AMRL-TR-75-61. Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratory, October 1975 (AD A-018074).

Pantle, A. J. and Sekuler, R. W. Velocity-sensitive elements in human vision: Initial psychophysical evidence. _Vision Research_, 1968, _8_, 445-450.

Petersik, J. T. Two types of apparent movement in the same visual display. Unpublished Master's Thesis, Miami University, 1975.

Sekuler, R. W., Pantle, A. J., and Levinson, E. Physiological basis of motion perception. In R. Held, H. Leibowitz, and H.-L. Teuber (Eds.), Handbook of sensory physiology, Volume VIII, Springer-Verlag, in press.

Stoper, A. E. and Banffy, S. Relation of split apparent motion to metacontrast. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3(2), 258-277.

☆U.S.Government Printing Office: 1977 — 757-080/283